Lecture Notes on

Mathematics for Economists

Chien-Fu CHOU

September 2006

Contents

1 Static Economic Models and The Concept of Equilibrium

Here we use three elementary examples to illustrate the general structure of an economic model.

1.1 Partial market equilibrium model

A partial market equilibrium model is constructed to explain the determination of the price of a certain commodity. The abstract form of the model is as follows.

$$Q_d = D(P; a) \quad Q_s = S(P; a) \quad Q_d = Q_s,$$

Q_d:	quantity demanded of the commodity	$D(P;a)$:	demand function
Q_s:	quantity supplied to the market	$S(P;a)$:	supply function
P:	market price of the commodity		
a:	a factor that affects demand and supply		

Equilibrium: A particular state that can be maintained.
Equilibrium conditions: Balance of forces prevailing in the model.
Substituting the demand and supply functions, we have $D(P;a) = S(P;a)$.
For a given a, we can solve this last equation to obtain the equilibrium price P^* as a function of a. Then we can study how a affects the market equilibrium price by inspecting the function.
Example: $D(P;a) = a^2/P$, $S(P) = 0.25P$. $a^2/P^* = 0.25P^* \Rightarrow P^* = 2a$, $Q_d^* = Q_s^* = 0.5a$.

1.2 General equilibrium model

Usually, markets for different commodities are interrelated. For example, the price of personal computers is strongly influenced by the situation in the market of microprocessors, the price of chicken meat is related to the supply of pork, etc. Therefore, we have to analyze interrelated markets within the same model to be able to capture such interrelationship and the partial equilibrium model is extended to the general equilibrium model. In microeconomics, we even attempt to include every commodity (including money) in a general equilibrium model.

$$\begin{array}{cccc} Q_{d1} = D_1(P_1, \ldots, P_n; a) & Q_{d2} = D_2(P_1, \ldots, P_n; a) & \ldots & Q_{dn} = D_n(P_1, \ldots, P_n; a) \\ Q_{s1} = S_1(P_1, \ldots, P_n; a) & Q_{s2} = S_2(P_1, \ldots, P_n; a) & & Q_{sn} = S_n(P_1, \ldots, P_n; a) \\ Q_{d1} = Q_{s1} & Q_{d2} = Q_{s2} & & Q_{dn} = Q_{sn} \end{array}$$

Q_{di}:	quantity demanded of commodity i
Q_{si}:	quantity supplied of commodity i
P_i:	market price of commodity i
a:	a factor that affects the economy
$D_i(P_1, \ldots, P_n; a)$:	demand function of commodity i
$S_i(P_1, \ldots, P_n; a)$:	supply function of commodity i

We have three variables and three equations for each commodity/market.

Substituting the demand and supply functions, we have

$$\begin{array}{rcccl} D_1(P_1,\ldots,P_n;a) - S_1(P_1,\ldots,P_n;a) & \equiv & E_1(P_1,\ldots,P_n;a) & = & 0 \\ D_2(P_1,\ldots,P_n;a) - S_2(P_1,\ldots,P_n;a) & \equiv & E_2(P_1,\ldots,P_n;a) & = & 0 \\ & \vdots & & & \vdots \\ D_n(P_1,\ldots,P_n;a) - S_n(P_1,\ldots,P_n;a) & \equiv & E_n(P_1,\ldots,P_n;a) & = & 0. \end{array}$$

For a given a, it is a simultaneous equation in $(P_1,\ldots,P_n)$. There are n equations and n unknown. In principle, we can solve the simultaneous equation to find the equilibrium prices $(P_1^*,\ldots,P_n^*)$.
A 2-market linear model:
$D_1 = a_0 + a_1P_1 + a_2P_2,\ S_1 = b_0 + b_1P_1 + b_2P_2,\ D_2 = \alpha_0 + \alpha_1P_1 + \alpha_2P_2,\ S_2 = \beta_0 + \beta_1P_1 + \beta_2P_2$.

$$(a_0 - b_0) + (a_1 - b_1)P_1 + (a_2 - b_2)P_2 = 0$$
$$(\alpha_0 - \beta_0) + (\alpha_1 - \beta_1)P_1 + (\alpha_2 - \beta_2)P_2 = 0.$$

1.3 National income model

The most fundamental issue in macroeconomics is the determination of the national income of a country.

$$\begin{array}{rcl} C & = & a + bY \ \ (a > 0,\ 0 < b < 1) \\ I & = & I(r) \\ Y & = & C + I + \bar{G} \\ S & = & Y - C. \end{array}$$

C:	Consumption	Y:	National income
I:	Investment	S:	Savings
$\bar{G}$:	government expenditure	r:	interest rate

a, b: coefficients of the consumption function.

To solve the model, we substitute the first two equations into the third to obtain $Y = a + bY + I_0 + \bar{G} \Rightarrow Y^* = (a + I(r) + \bar{G})/(1 - b)$.

1.4 The ingredients of a model

We set up economic models to study economic phenomena (cause-effect relationships), or how certain economic variables affect other variables. A model consists of equations, which are relationships among variables.
Variables can be divided into three categories:
Endogenous variables: variables we choose to represent different states of a model.
Exogenous variables: variables assumed to affect the endogenous variables but are not affected by them.
Causes (Changes in exogenous var.) $\Rightarrow$ Effects (Changes in endogenous var.)
Parameters: Coefficients of the equations.

	End. Var.	Ex. Var.	Parameters
Partial equilibrium model:	P, Q_d, Q_s	a	Coefficients of $D(P;a)$, $S(P;a)$
General equilibrium model:	P_i, Q_{di}, Q_{si}	a	
Income model:	C, Y, I, S	r, $\bar{G}$	a, b

Equations can be divided into three types:
Behavioral equations: representing the decisions of economic agents in the model.
Equilibrium conditions: the condition such that the state can be maintained (when different forces/motivations are in balance).
Definitions: to introduce new variables into the model.

	Behavioral equations	Equilibrium cond.	Definitions
Partial equilibrium model:	$Q_d = D(P;a)$, $Q_s = S(P;a)$	$Q_d = Q_s$	
General equilibrium model:	$Q_{di} = D_i(P_1, \ldots, P_n)$, $Q_{si} = S_i(P_1, \ldots, P_n)$	$Q_{di} = Q_{si}$	
Income model:	$C = a + bY$, $I = I(r)$	$Y = C + I + \bar{G}$	$S = Y - C$

1.5 The general economic model

Assume that there are n endogenous variables and m exogenous variables.
Endogenous variables: $x_1, x_2, \ldots, x_n$
Exogenous variables: $y_1, y_2, \ldots, y_m$.
There should be n equations so that the model can be solved.

$$\begin{aligned} F_1(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) &= 0 \\ F_2(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) &= 0 \\ &\vdots \\ F_n(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) &= 0. \end{aligned}$$

Some of the equations are behavioral, some are equilibrium conditions, and some are definitions.

In principle, given the values of the exogenous variables, we solve to find the endogenous variables as functions of the exogenous variables:

$$\begin{aligned} x_1 &= x_1(y_1, y_2, \ldots, y_m) \\ x_2 &= x_1(y_1, y_2, \ldots, y_m) \\ &\vdots \\ x_n &= x_1(y_1, y_2, \ldots, y_m). \end{aligned}$$

If the equations are all linear in $(x_1, x_2, \ldots, x_n)$, then we can use Cramer's rule (to be discussed in the next part) to solve the equations. However, if some equations are nonlinear, it is usually very difficult to solve the model. In general, we use comparative statics method (to be discussed in part 3) to find the differential relationships between x_i and y_j: $\frac{\partial x_i}{\partial y_j}$.

1.6 Problems

1. Find the equilibrium solution of the following model:

$$Q_d = 3 - P^2, \qquad Q_s = 6P - 4, \qquad Q_s = Q_d.$$

2. The demand and supply functions of a two-commodity model are as follows:

$$\begin{aligned} Q_{d1} &= 18 - 3P_1 + P_2 & Q_{d2} &= 12 + P_1 - 2P_2 \\ Q_{s1} &= -2 + 4P_1 & Q_{s2} &= -2 + 3P_2 \end{aligned}$$

   Find the equilibrium of the model.

3. (The effect of a sales tax) Suppose that the government imposes a sales tax of t dollars per unit on product 1. The partial market model becomes

$$Q_1^d = D(P_1 + t), \quad Q_1^s = S(P_1), \quad Q_1^d = Q_1^s.$$

   Eliminating Q_1^d and Q_1^s, the equilibrium price is determined by $D(P_1 + t) = S(P_1)$.

   (a) Identify the endogenous variables and exogenous variable(s).

   (b) Let $D(p) = 120 - P$ and $S(p) = 2P$. Calculate P_1 and Q_1 both as function of t.

   (c) If t increases, will P_1 and Q_1 increase or decrease?

4. Let the national-income model be:

$$\begin{aligned} Y &= C + I_0 + G \\ C &= a + b(Y - T_0) & (a > 0, 0 < b < 1) \\ G &= gY & (0 < g < 1) \end{aligned}$$

   (a) Identify the endogenous variables.

   (b) Give the economic meaning of the parameter g.

   (c) Find the equilibrium national income.

   (d) What restriction(s) on the parameters is needed for an economically reasonable solution to exist?

5. Find the equilibrium Y and C from the following:

$$Y = C + I_0 + G_0, \qquad C = 25 + 6Y^{1/2}, \qquad I_0 = 16, \qquad G_0 = 14.$$

6. In a 2-good market equilibrium model, the **inverse** demand functions are given by

$$P_1 = Q_1^{\frac{-2}{3}} Q_2^{\frac{1}{3}}, \quad P_2 = Q_1^{\frac{1}{3}} Q_2^{\frac{-2}{3}}.$$

   (a) Find the demand functions $Q_1 = D^1(P_1, P_2)$ and $Q_2 = D^2(P_1, P_2)$.

   (b) Suppose that the supply functions are

$$Q_1 = a^{-1} P_1, \qquad Q_2 = P_2.$$

   Find the equilibrium prices (P_1^*, P_2^*) and quantities (Q_1^*, Q_2^*) as functions of a.

2 Matrix Algebra

A matrix is a two dimensional rectangular array of numbers:

$$A \equiv \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

There are n columns each with m elements or m rows each with n elements. We say that the size of A is $m \times n$.
If $m = n$, then A is a square matrix.
A $m \times 1$ matrix is called a column vector and a $1 \times n$ matrix is called a row vector.
A 1×1 matrix is just a number, called a scalar number.

2.1 Matrix operations

Equality: $A = B \Rightarrow$ (1) size(A) = size(B), (2) $a_{ij} = b_{ij}$ for all ij.
Addition/subtraction: $A+B$ and $A-B$ can be defined only when size(A) = size(B), in that case, size($A+B$) = size($A-B$) = size(A) = size(B) and $(A+B)_{ij} = a_{ij} + b_{ij}$, $(A-B)_{ij} = a_{ij} - b_{ij}$. For example,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow A + B = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix}, \quad A - B = \begin{pmatrix} 0 & 2 \\ 3 & 3 \end{pmatrix}.$$

Scalar multiplication: The multiplication of a scalar number α and a matrix A, denoted by αA, is always defined with size (αA) = size(A) and $(\alpha A)_{ij} = \alpha a_{ij}$. For example, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \Rightarrow 4A = \begin{pmatrix} 4 & 8 \\ 12 & 16 \end{pmatrix}$.

Multiplication of two matrices: Let size(A) $= m \times n$ and size(B) $= o \times p$, the multiplication of A and B, $C = AB$, is more complicated. (1) it is not always defined. (2) $AB \neq BA$ even when both are defined. The condition for AB to be meaningful is that the number of columns of A should be equal to the number of rows of B, i.e., $n = o$. In that case, size (AB) = size $(C) = m \times p$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & a_{np} \end{pmatrix} \Rightarrow C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{pmatrix},$$

where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.
Examples: $\begin{pmatrix} 1 & 2 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3+8 \\ 0+20 \end{pmatrix} = \begin{pmatrix} 11 \\ 20 \end{pmatrix}$,
$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 5+14 & 6+16 \\ 15+28 & 18+32 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$,

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5+18 & 10+24 \\ 7+24 & 14+32 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

Notice that $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \neq \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

2.2 Matrix representation of a linear simultaneous equation system

A linear simultaneous equation system:

$$\begin{array}{ccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n = b_1 \\ & \vdots & & \vdots & \\ a_{n1}x_1 & + & \dots & + & a_{nn}x_n = b_n \end{array}$$

Define $A \equiv \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, $x \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $b \equiv \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Then the equation $Ax = b$ is equivalent to the simultaneous equation system.

Linear 2-market model:

$$E_1 = (a_1 - b_1)p_1 + (a_2 - b_2)p_2 + (a_0 - b_0) = 0$$
$$E_2 = (\alpha_1 - \beta_1)p_1 + (\alpha_2 - \beta_2)p_2 + (\alpha_0 - \beta_0) = 0$$

$$\Rightarrow \begin{pmatrix} a_1 - b_1 & a_2 - b_2 \\ \alpha_1 - \beta_1 & \alpha_2 - \beta_2 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} a_0 - b_0 \\ \alpha_0 - \beta_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Income determination model:

$$\begin{array}{c} C = a + bY \\ I = I(r) \\ Y = C + I \end{array} \quad \Rightarrow \begin{pmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} C \\ I \\ Y \end{pmatrix} = \begin{pmatrix} a \\ I(r) \\ 0 \end{pmatrix}.$$

In the algebra of real numbers, the solution to the equation $ax = b$ is $x = a^{-1}b$. In matrix algebra, we wish to define a concept of A^{-1} for a $n \times n$ matrix A so that $x = A^{-1}b$ is the solution to the equation $Ax = b$.

2.3 Commutative, association, and distributive laws

The notations for some important sets are given by the following table.

N = nature numbers $1, 2, 3, \dots$	I = integers $\dots, -2, -1, 0, 1, 2, \dots$
Q = rational numbers $\frac{m}{n}$	R = real numbers
R^n = n-dimensional column vectors	$\mathcal{M}(m, n) = m \times n$ matrices
$\mathcal{M}(n) = n \times n$ matrices	

A binary operation is a law of composition of two elements from a set to form a third element of the same set. For example, $+$ and $\times$ are binary operations of real numbers R.

Another important example: *addition* and *multiplication* are binary operations of matrices.

Commutative law of $+$ and $\times$ in R: $a+b=b+a$ and $a\times b=b\times a$ for all $a,b\in R$.
Association law of $+$ and $\times$ in R: $(a+b)+c=a+(b+c)$ and $(a\times b)\times c=a\times(b\times c)$ for all $a,b,c\in R$.
Distributive law of $+$ and $\times$ in R: $(a+b)\times c=a\times c+b\times c$ and $c\times(a+b)=c\times a+c\times b$ for all $a,b,c\in R$.
The *addition* of matrices satisfies both commutative and associative laws: $A+B=B+A$ and $(A+B)+C=A+(B+C)$ for all $A,B,C\in\mathcal{M}(m,n)$. The proof is trivial.
In an example, we already showed that the matrix multiplication does not satisfy the commutative law $AB\neq BA$ even when both are meaningful.
Nevertheless the matrix multiplication satisfies the associative law $(AB)C=A(BC)$ when the sizes are such that the multiplications are meaningful. However, this deserves a proof!
It is also true that matrix *addition* and *multiplication* satisfy the distributive law: $(A+B)C=AC+BC$ and $C(A+B)=CA+CB$. You should try to prove these statements as exercises.

2.4 Special matrices

In the space of real numbers, 0 and 1 are very special. 0 is the unit element of $+$ and 1 is the unit element of $\times$: $0+a=a+0=a$, $0\times a=a\times 0=0$, and $1\times a=a\times 1=a$. In matrix algebra, we define zero matrices and identity matrices as

$$O_{m,n}\equiv\begin{pmatrix}0&\dots&0\\ \vdots&\ddots&\vdots\\ 0&\dots&0\end{pmatrix}\quad I_n\equiv\begin{pmatrix}1&0&\dots&0\\ 0&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&1\end{pmatrix}.$$

Clearly, $O+A=A+O=A$, $OA=AO=O$, and $IA=AI=A$. In the multiplication of real numbers if $a,b\neq 0$ then $a\times b\neq 0$. However, $\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=O_{2,2}$.
Idempotent matrix: If $AA=A$ (A must be square), then A is an idempotent matrix. Both $O_{n,n}$ and I_n are idempotent. Another example is $A=\begin{pmatrix}0.5&0.5\\0.5&0.5\end{pmatrix}$.
Transpose of a matrix: For a matrix A with size $m\times n$, we define its transpose A' as a matrix with size $n\times m$ such that the ij-th element of A' is equal to the ji-th element of A, $a'_{ij}=a_{ji}$.

$$A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\ \text{then}\ A'=\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}.$$

Properties of matrix transposition:
(1) $(A')'=A$, (2) $(A+B)'=A'+B'$, (3) $(AB)'=B'A'$.
Symmetrical matrix: If $A=A'$ (A must be square), then A is symmetrical. The

condition for A to be symmetrical is that $a_{ij} = a_{ji}$. Both $O_{n,n}$ and I_n are symmetrical. Another example is $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$.
Projection matrix: A symmetrical idempotent matrix is a projection matrix.
Diagonal matrix: A symmetrical matrix A is diagonal if $a_{ij} = 0$ for all $i \neq j$. Both I_n and $O_{n,n}$ are diagonal. Another example is $A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

2.5 Inverse of a square matrix

We are going to define the inverse of a square matrix $A \in \mathcal{M}(n)$.
Scalar: $aa^{-1} = a^{-1}a = 1 \Rightarrow$ if b satisfies $ab = ba = 1$ then $b = a^{-1}$.
Definition of A^{-1}: If there exists a $B \in \mathcal{M}(n)$ such that $AB = BA = I_n$, then we define $A^{-1} = B$.
Examples: (1) Since $II = I$, $I^{-1} = I$. (2) $O_{n,n}B = O_{n,n} \Rightarrow O_{n,n}^{-1}$ does not exist. (3) If $A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$, $a_1, a_2 \neq 0$, then $A^{-1} = \begin{pmatrix} a_1^{-1} & 0 \\ 0 & a_2^{-1} \end{pmatrix}$. (4) If $a_1 = 0$ or $a_2 = 0$, then A^{-1} does not exist.
Singular matrix: A square matrix whose inverse matrix does not exist.
Non-singular matrix: A is non-singular if A^{-1} exists.

Properties of matrix inversion:
Let $A, B \in \mathcal{M}(n)$, (1) $(A^{-1})^{-1} = A$, (2) $(AB)^{-1} = B^{-1}A^{-1}$, (3) $(A')^{-1} = (A^{-1})'$.

2.6 Problems

1. Let $A = I - X(X'X)^{-1}X'$.
   (a) If the dimension of X is $m \times n$, what must be the dimension of I and A.
   (b) Show that matrix A is idempotent.

2. Let A and B be $n \times n$ matrices and I be the identity matrix.
   (a) $(A + B)^3 = ?$
   (b) $(A + I)^3 = ?$

3. Let $B = \left(\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}\right)$, $U = (1, 1)'$, $V = (1, -1)'$, and $W = aU + bV$, where a and b are real numbers. Find BU, BV, and BW. Is B idempotent?

4. Suppose A is a $n \times n$ nonsingular matrix and P is a $n \times n$ idempotent matrix. Show that APA^{-1} is idempotent.

5. Suppose that A and B are $n \times n$ symmetric idempotent matrices and $AB = B$. Show that $A - B$ is idempotent.

6. Calculate $(x_1, x_2)\begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

7. Let $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(a) Show that $J^2 = -I$.

(b) Make use of the above result to calculate J^3, J^4, and J^{-1}.

(c) Show that $(aI + bJ)(cI + dJ) = (ac - bd)I + (ad + bc)J$.

(d) Show that $(aI + bJ)^{-1} = \dfrac{1}{a^2 + b^2}(aI - bJ)$ and $[(\cos\theta)I + (\sin\theta)J]^{-1} = (\cos\theta)I - (\sin\theta)J$.

3 Vector Space and Linear Transformation

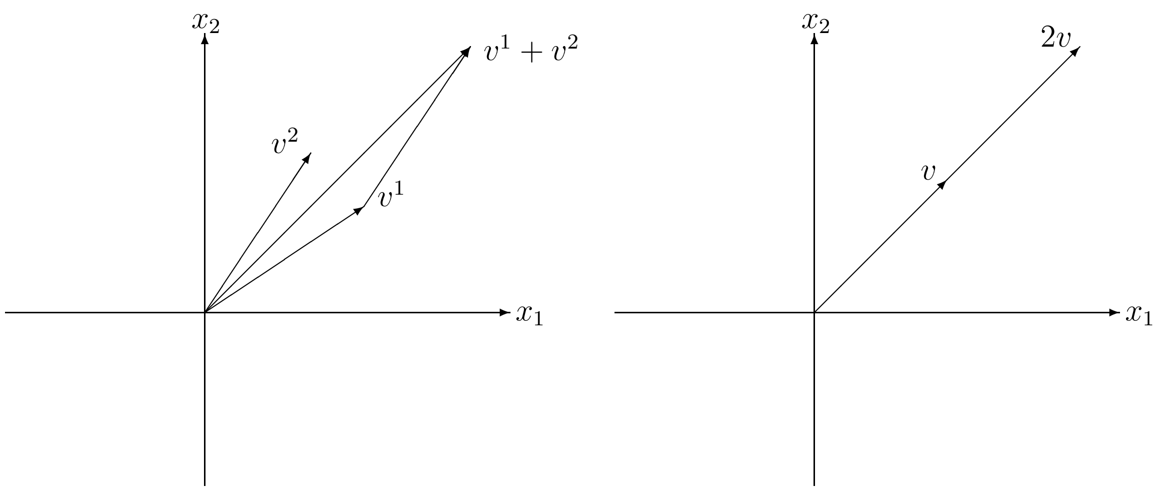
In the last section, we regard a matrix simply as an array of numbers. Now we are going to provide some geometrical meanings to a matrix.
(1) A matrix as a collection of column (row) vectors
(2) A matrix as a linear transformation from a vector space to another vector space

3.1 Vector space, linear combination, and linear independence

Each point in the m-dimensional Euclidean space can be represented as a m-dimensional column vector $v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$, where each v_i represents the i-th coordinate. Two points in the m-dimensional Euclidean space can be added according to the rule of matrix addition. A point can be multiplied by a scalar according to the rule of scalar multiplication.

Vector addition: $\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_m + w_m \end{pmatrix}$.

Scalar multiplication: $\alpha \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_m \end{pmatrix}$.



With such a structure, we say that the m-dimensional Euclidean space is a vector space.

m-dimensional column vector space: $R^m = \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}, \; v_i \in R \right\}$.

We use superscripts to represent individual vectors.

A $m \times n$ matrix: a collection of n m-dimensional column vectors:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \left\{ \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\}$$

Linear combination of a collection of vectors $\{v^1, \dots, v^n\}$: $w = \sum_{i=1}^{n} \alpha_i v^i$, where $(\alpha_1, \dots, \alpha_n) \neq (0, \dots, 0)$.

Linear dependence of $\{v^1, \dots, v^n\}$: If one of the vectors is a linear combination of others, then the collection is said to be linear dependent. Alternatively, the collection is linearly dependent if $(0, \dots, 0)$ is a linear combination of it.

Linear independence of $\{v^1, \dots, v^n\}$: If the collection is not linear dependent, then it is linear independent.

Example 1: $v^1 = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}, v^2 = \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix}, v^3 = \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix}, \; a_1 a_2 a_3 \neq 0$.
If $\alpha_1 v^1 + \alpha_2 v^2 + \alpha_3 v^3 = 0$ then $(\alpha_1, \alpha_2, \alpha_3) = (0, 0, 0)$. Therefore, $\{v^1, v^2, v^3\}$ must be linear independent.
Example 2: $v^1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v^2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, v^3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$.
$2v^2 = v^1 + v^3$. Therefore, $\{v^1, v^2, v^3\}$ is linear dependent.
Example 3: $v^1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v^2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.
$\alpha_1 v^1 + \alpha_2 v^2 = \begin{pmatrix} \alpha_1 + 4\alpha_2 \\ 2\alpha_1 + 5\alpha_2 \\ 3\alpha_1 + 6\alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \alpha_1 = \alpha_2 = 0$. Therefore, $\{v^1, v^2\}$ is linear independent.

Span of $\{v^1, \dots, v^n\}$: The space of linear combinations.
If a vector is a linear combination of other vectors, then it can be removed without changing the span.

Rank$\begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{m1} & \dots & v_{mn} \end{pmatrix}$ $\equiv$ Dimension(Span$\{v^1, \dots, v^n\}$) = Maximum # of independent vectors.

3.2 Linear transformation

Consider a $m \times n$ matrix A. Given $x \in R^n$, $Ax \in R^m$. Therefore, we can define a **linear transformation** from R^n to R^m as $f(x) = Ax$ or

$$f : R^n \rightarrow R^m, \quad f(x) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

It is linear because $f(\alpha x + \beta w) = A(\alpha x + \beta w) = \alpha Ax + \beta Aw = \alpha f(x) + \beta f(w)$.

Standard basis vectors of R^n: $e^1 \equiv \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e^2 \equiv \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e^n \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.

Let v^i be the i-th column of A, $v^i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$.

$v^i = f(e^i)$: $\begin{pmatrix} a_{11} \\ \dots \\ a_{m1} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow v^1 = f(e^1) = Ae^1$, etc.

Therefore, v^i is the image of the i-th standard basis vector e^i under f.
Span$\{v^1, \dots, v^n\}$ = Range space of $f(x) = Ax \equiv R(A)$.
Rank$(A) \equiv \dim(R(A))$.
Null space of $f(x) = Ax$: $N(A) \equiv \{x \in R^n, f(x) = Ax = 0\}$.
$\dim(R(A)) + \dim(N(A)) = n$.

Example 1: $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. $N(A) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$, $R(A) = R^2$, Rank$(A) = 2$.
Example 2: $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. $N(B) = \left\{ \begin{pmatrix} k \\ -k \end{pmatrix}, k \in R \right\}$, $R(B) = \left\{ \begin{pmatrix} k \\ k \end{pmatrix}, k \in R \right\}$,
Rank$(B) = 1$.

The multiplication of two matrices can be interpreted as the composition of two linear transformations.

$$f: \ R^n \rightarrow R^m, \ f(x) = Ax, \quad g: \ R^p \rightarrow R^n, \ g(y) = By, \quad \Rightarrow f(g(x)) = A(By), \ f \circ g: \ R^p \rightarrow R^m.$$

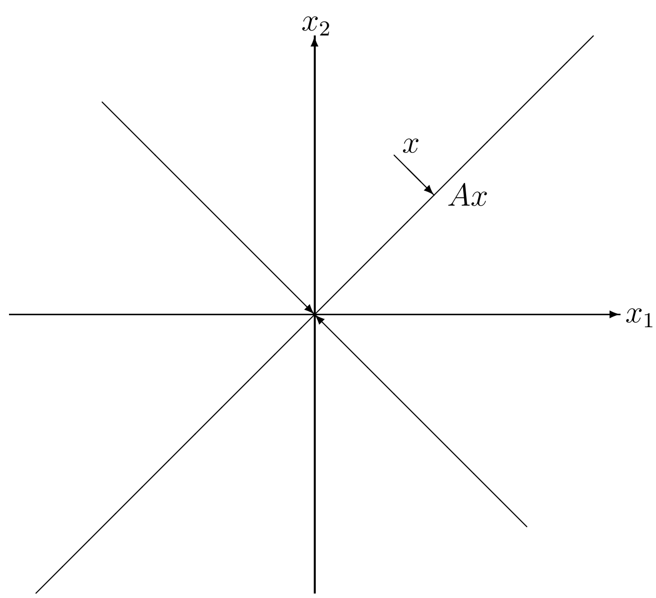
The composition is meaningful only when the dimension of the range space of $g(y)$ is equal to the dimension of the domain of $f(x)$, which is the same condition for the validity of the matrix multiplication.

Every linear transformation $f : R^n \rightarrow R^m$ can be represented by $f(x) = Ax$ for some $m \times n$ matrix.

3.3 Inverse transformation and inverse of a square matrix

Consider now the special case of square matrices. Each $A \in \mathcal{M}(n)$ represents a linear transformation $f : R^n \rightarrow R^n$.
The definition of the inverse matrix A^{-1} is such that $AA^{-1} = A^{-1}A = I$. If we regard A as a linear transformation from $R^n \rightarrow R^n$ and I as the identity transformation that maps every vector (point) into itself, then A^{-1} is the inverse mapping of A. If $\dim(N(A)) = 0$, then $f(x)$ is one to one.
If $\dim(R(A)) = n$, then $R(A) = R^n$ and $f(x)$ is onto.
$\Rightarrow$ if Rank$(A) = n$, then $f(x)$ is one to one and onto and there exits an inverse mapping $f^{-1} : R^n \rightarrow R^n$ represented by a $n \times n$ square matrix A^{-1}. $f^{-1}f(x) = x \Rightarrow A^{-1}Ax = x$.
$\Rightarrow$ if Rank$(A) = n$, then A is non-singular.
if Rank$(A) < n$, then $f(x)$ is not onto, no inverse mapping exists, and A is singular.
Examples: Rank $\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} = 3$ and Rank $\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} = 2$.

Remark: $O_{n,n}$ represents the mapping that maps every point to the origin. I_n represents the identity mapping that maps a point to itself. A projection matrix represents a mapping that projects points onto a linear subspace of R^n, eg., $\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$ projects points onto the 45 degree line.



$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$
$$Ax = \begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix}$$
$$x' = \begin{pmatrix} k \\ -k \end{pmatrix}, \; Ax' = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

3.4 Problems

1. Let $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

   and T_B the corresponding linear transformation $T_B : R^3 \to R^3$, $T_B(x) = Bx$, where $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in R^3$.

   (a) Is $v^1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$, $a \neq 0$, in the null space of T_B? Why or why not?

   (b) Is $v^2 = \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix}$, $b \neq 0$, in the range space of T_B? Why or why not? How about $v^3 = \begin{pmatrix} c \\ d \\ 0 \end{pmatrix}$?

   (c) Find Rank(B).

2. Let A be an idempotent matrix.

   (a) Show that $I - A$ is also idempotent.

   (b) Suppose that $x \neq 0$ is in the null space of A, i.e., $Ax = 0$. Show that x must be in the range space of $I - A$, i.e., show that there exists a vector y such that $(I - A)y = x$. (Hint: Try $y = x$.)

   (c) Suppose that y is in the range space of A. Show that y must be in the null space of $I - A$.

   (d) Suppose that A is $n \times n$ and Rank$[A] = n - k$, $n > k > 0$. What is the rank of $I - A$?

3. Let $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, $x = \begin{pmatrix} 1 \\ a \\ b \end{pmatrix}$, $y = \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}$, and $B = I - A$.

   (a) Calculate AA and BB.

   (b) If y is in the range space of A, what are the values of α and β?

   (c) What is the dimension of the range space of A?

   (d) Determine the rank of A.

   (e) Suppose now that x is in the null space of B. What should be the values of a and b?

   (f) What is the dimension of the null space of B?

(g) Determine the rank of B?

4. Let $A = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

   (a) Determine the ranks of A and B.

   (b) Determine the null space and range space of each of A and B and explain why.

   (c) Determine whether they are idempotent.

4 Determinant, Inverse Matrix, and Cramer's rule

In this section we are going to derive a general method to calculate the inverse of a square matrix. First, we define the determinant of a square matrix. Using the properties of determinants, we find a procedure to compute the inverse matrix. Then we derive a general procedure to solve a simultaneous equation.

4.1 Permutation group

A permutation of $\{1, 2, \ldots, n\}$ is a 1-1 mapping of $\{1, 2, \ldots, n\}$ onto itself, written as $\pi = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}$ meaning that 1 is mapped to i_1, 2 is mapped to i_2, ..., and n is mapped to i_n. We also write $\pi = (i_1, i_2, \ldots, \imath_n)$ when no confusing.

Permutation set of $\{1, 2, \ldots, n\}$: $\mathcal{P}_n \equiv \{\pi = (i_1, i_2, \ldots, i_n) : \quad \pi \text{ is a permutation}\}$.

$\mathcal{P}_2 = \{(1, 2), (2, 1)\}$.

$\mathcal{P}_3 = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$.

$\mathcal{P}_4$: $4! = 24$ permutations.

Inversions in a permutation $\pi = (i_1, i_2, \ldots, i_n)$: If there exist k and l such that $k < l$ and $i_k > i_l$, then we say that an inversion occurs.

$N(i_1, i_2, \ldots, \imath_n)$: Total number of inversions in $(i_1, i_2, \ldots, i_n)$.

Examples: 1. $N(1, 2) = 0$, $N(2, 1) = 1$.

2. $N(1, 2, 3) = 0$, $N(1, 3, 2) = 1$, $N(2, 1, 3) = 1$,
$N(2, 3, 1) = 2$, $N(3, 1, 2) = 2$, $N(3, 2, 1) = 3$.

4.2 Determinant

Determinant of $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$:

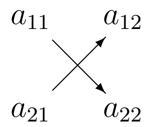
$$|A| \equiv \sum_{(i_1, i_2, \ldots, i_n) \in \mathcal{P}_n} (-1)^{N(i_1, i_2, \ldots, i_n)} a_{1i_1} a_{2i_2} \ldots a_{ni_n}.$$

$n = 2$: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (-1)^{N(1,2)} a_{11} a_{22} + (-1)^{N(2,1)} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}$.

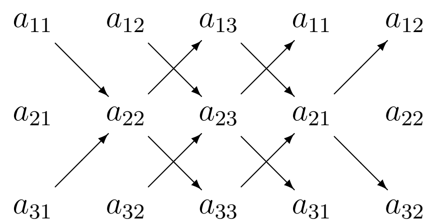
$n = 3$: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$

$(-1)^{N(1,2,3)} a_{11} a_{22} a_{33} + (-1)^{N(1,3,2)} a_{11} a_{23} a_{32} + (-1)^{N(2,1,3)} a_{12} a_{21} a_{33} +$
$(-1)^{N(2,3,1)} a_{12} a_{23} a_{31} + (-1)^{N(3,1,2)} a_{13} a_{21} a_{32} + (-1)^{N(3,2,1)} a_{13} a_{22} a_{31}$
$= a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31}$.

$n = 2$:



$n = 3$:



4.3 Properties of determinant

Property 1: $|A'| = |A|$.
Proof: Each term of $|A'|$ corresponds to a term of $|A|$ of the same sign.
By property 1, we can replace "column vectors" in the properties below by "row vectors".
Since a $n \times n$ matrix can be regarded as n column vectors $A = \{v^1, v^2, \ldots, v^n\}$, we can regard determinants as a function of n column vectors $|A| = D(v^1, v^2, \ldots, v^n)$, $D: R^{n\times n} \to R$.
By property 1, we can replace "column vectors" in the properties below by "row vectors".

Property 2: If two column vectors are interchanged, the determinant changes sign.
Proof: Each term of the new determinant corresponds to a term of $|A|$ of opposite sign because the number of inversion increases or decreases by 1.
Example: $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \times 4 - 2 \times 3 = -2$, $\begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix} = 2 \times 3 - 1 \times 4 = 2$,

Property 3: If two column vectors are identical, then the determinant is 0.
Proof: By property 2, the determinant is equal to the negative of itself, which is possible only when the determinant is 0.

Property 4: If you add a linear combination of other column vectors to a column vector, the determinant does not change.
Proof: Given other column vectors, the determinant function is a linear function of v^i: $D(\alpha v^i + \beta w^i;\ \text{other vectors}) = \alpha D(v^i;\ \text{other vectors}) + \beta D(w^i;\ \text{other vectors})$.
Example: $\begin{vmatrix} 1+5\times 2 & 2 \\ 3+5\times 4 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} + \begin{vmatrix} 5\times 2 & 2 \\ 5\times 4 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} + 5 \times \begin{vmatrix} 2 & 2 \\ 4 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} + 5 \times 0$.

Submatrix: We denote by A_{ij} the submatrix of A obtained by deleting the i-th row and j-th column from A.
Minors: The determinant $|A_{ij}|$ is called the minor of the element a_{ij}.
Cofactors: $C_{ij} \equiv (-1)^{i+j}|A_{ij}|$ is called the cofactor of a_{ij}.

Property 5 (Laplace theorem): Given $i = \bar{i}$, $|A| = \sum_{j=1}^{n} a_{\bar{i}j} C_{\bar{i}j}$.

Given $j = \bar{j}$, $|A| = \sum_{i=1}^{n} a_{i\bar{j}} C_{i\bar{j}}$.
Proof: In the definition of the determinant of $|A|$, all terms with a_{ij} can be put togethere to become $a_{ij}C_{ij}$.

Example: $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{vmatrix} = 1 \times \begin{vmatrix} 5 & 6 \\ 8 & 0 \end{vmatrix} - 2 \times \begin{vmatrix} 4 & 6 \\ 7 & 0 \end{vmatrix} + 3 \times \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}$.

Property 6: Given $i' \neq \bar{i}$, $\sum_{j=1}^{n} a_{i'j} C_{\bar{i}j} = 0$.

Given $j' \neq \bar{j}$, $= \sum_{i=1}^{n} a_{ij'} C_{i\bar{j}} = 0$.
Therefore, if you multiply cofactors by the elements from a different row or column, you get 0 instead of the determinant.
Proof: The sum becomes the determinant of a matrix with two identical rows (columns).

Example: $0 = 4 \times \begin{vmatrix} 5 & 6 \\ 8 & 0 \end{vmatrix} - 5 \times \begin{vmatrix} 4 & 6 \\ 7 & 0 \end{vmatrix} + 6 \times \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}$.

4.4 Computation of the inverse matrix

Using properties 5 and 6, we can calculate the inverse of A as follows.

1. Cofactor matrix: $C \equiv \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{pmatrix}$.

2. Adjoint of A: Adj $A \equiv C' = \begin{pmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{pmatrix}$.

3. $\Rightarrow AC' = C'A = \begin{pmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |A| \end{pmatrix} \Rightarrow$ if $|A| \neq 0$ then $\frac{1}{|A|}C' = A^{-1}$.

Example 1: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ then $C = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$.

$$A^{-1} = \frac{1}{|A|}C' = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}; \text{ if } |A| = a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

Example 2: $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \Rightarrow |A| = 27 \neq 0$ and

$C_{11} = \begin{vmatrix} 5 & 6 \\ 8 & 0 \end{vmatrix} = -48$, $C_{12} = -\begin{vmatrix} 4 & 6 \\ 7 & 0 \end{vmatrix} = 42$, $C_{13} = \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3$,

$C_{21} = -\begin{vmatrix} 2 & 3 \\ 8 & 0 \end{vmatrix} = 24$, $C_{22} = \begin{vmatrix} 1 & 3 \\ 7 & 0 \end{vmatrix} = -21$, $C_{23} = -\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = 6$,
$C_{31} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3$, $C_{32} = -\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 6$, $C_{33} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3$,

$$C = \begin{pmatrix} -48 & 42 & -3 \\ 24 & -21 & 6 \\ -3 & 6 & -3 \end{pmatrix}, \quad C' = \begin{pmatrix} -48 & 24 & -3 \\ 42 & -21 & 6 \\ -3 & 6 & -3 \end{pmatrix}, \quad A^{-1} = \frac{1}{27}\begin{pmatrix} -48 & 24 & -3 \\ 42 & -21 & 6 \\ -3 & 6 & -3 \end{pmatrix}.$$

If $|A| = 0$, then A is singular and A^{-1} does not exist. The reason is that $|A| = 0 \Rightarrow AC' = 0_{n\times n} \Rightarrow C_{11}\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \cdots + C_{n1}\begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. The column vectors of A are linear dependent, the linear transformation T_A is not onto and therefore an inverse transformation does not exist.

4.5 Cramer's rule

If $|A| \neq 0$ then A is non-singular and $A^{-1} = \dfrac{C'}{|A|}$. The solution to the simultaneous equation $Ax = b$ is $x = A^{-1}b = \dfrac{C'b}{|A|}$.

Cramer's rule: $x_i = \dfrac{\sum_{j=1}^{n} C_j b_j}{|A|} = \dfrac{|A_i|}{|A|}$, where A_i is a matrix obtained by replacing the i-th column of A by b, $A_i = \{v^1, \ldots, v^{i-1}, b, v^{i+1}, \ldots, v^n\}$.

4.6 Economic applications

Linear 2-market model: $\begin{pmatrix} a_1 - b_1 & a_2 - b_2 \\ \alpha_1 - \beta_1 & \alpha_2 - \beta_2 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} b_0 - a_0 \\ \beta_0 - \alpha_0 \end{pmatrix}$.

$$p_1 = \frac{\begin{vmatrix} b_0 - a_0 & a_2 - b_2 \\ \beta_0 - \alpha_0 & \alpha_2 - \beta_2 \end{vmatrix}}{\begin{vmatrix} a_1 - b_1 & a_2 - b_2 \\ \alpha_1 - \beta_1 & \alpha_2 - \beta_2 \end{vmatrix}}, \quad p_2 = \frac{\begin{vmatrix} a_1 - b_1 & b_0 - a_0 \\ \alpha_1 - \beta_1 & \beta_0 - \alpha_0 \end{vmatrix}}{\begin{vmatrix} a_1 - b_1 & a_2 - b_2 \\ \alpha_1 - \beta_1 & \alpha_2 - \beta_2 \end{vmatrix}}.$$

Income determination model: $\begin{pmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} C \\ I \\ Y \end{pmatrix} = \begin{pmatrix} a \\ I(r) \\ 0 \end{pmatrix}$.

$$C = \frac{\begin{vmatrix} a & 0 & -b \\ I(r) & 1 & 0 \\ 0 & 1 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{vmatrix}}, \quad I = \frac{\begin{vmatrix} 1 & a & -b \\ 0 & I(r) & 0 \\ 1 & 0 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{vmatrix}}, \quad Y = \frac{\begin{vmatrix} 1 & 0 & a \\ 0 & 1 & I(r) \\ 1 & 1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{vmatrix}}.$$

IS-LM model: In the income determination model, we regard interest rate as given and consider only the product market. Now we enlarge the model to include the money market and regard interest rate as the price (an endogenous variable) determined in the money market.

good market:	money market:
$C = a + bY$	$L = kY - lR$
$I = I_0 - iR$	$M = \bar{M}$
$C + I + \bar{G} = Y$	$M = L$

end. var: C, I, Y, R (interest rate), L(demand for money), M (money supply)
ex. var: $\bar{G}$, $\bar{M}$ (quantity of money). parameters: a, b, i, k, l.
Substitute into equilibrium conditions:

good market:	money market	endogenous variables:
$a + bY + I_0 - iR + \bar{G} = Y,$	$kY - lR = \bar{M},$	Y, R

$$\begin{pmatrix} 1-b & i \\ k & -l \end{pmatrix} \begin{pmatrix} Y \\ R \end{pmatrix} = \begin{pmatrix} a + I_0 + \bar{G} \\ \bar{M} \end{pmatrix}$$

$$Y = \frac{\begin{vmatrix} a + I_0 + \bar{G} & i \\ \bar{M} & -l \end{vmatrix}}{\begin{vmatrix} 1-b & i \\ k & -l \end{vmatrix}}, \quad R = \frac{\begin{vmatrix} 1-b & a + I_0 + \bar{G} \\ k & \bar{M} \end{vmatrix}}{\begin{vmatrix} 1-b & i \\ k & -l \end{vmatrix}}.$$

Two-country income determination model: Another extension of the income determination model is to consider the interaction between domestic country and the rest of the world (foreign country).

domestic good market:	foreign good market:	endogenous variables:
$C = a + bY$	$C' = a' + b'Y'$	C, I, Y,
$I = I_0$	$I' = I'_0$	M (import),
$M = M_0 + mY$	$M' = M'_0 + m'Y'$	X (export),
$C + I + X - M = Y$	$C' + I' + X' - M' = Y'$	C', I', Y', M', X'.

By definition, $X = M'$ and $X' = M$. Substituting into the equilibrium conditions,

$$(1 - b + m)Y - m'Y' = a + I_0 + M'_0 - M_0 \quad (1 - b' + m')Y' - mY = a' + I'_0 + M_0 - M'_0.$$

$$\begin{pmatrix} 1-b+m & -m' \\ -m & 1-b'+m' \end{pmatrix} \begin{pmatrix} Y \\ Y' \end{pmatrix} = \begin{pmatrix} a + I_0 + M'_0 - M_0 \\ a' + I'_0 + M_0 - M'_0 \end{pmatrix}.$$

$$Y = \frac{\begin{vmatrix} a + I_0 + M'_0 - M_0 & -m' \\ a' + I'_0 + M_0 - M'_0 & 1-b'+m' \end{vmatrix}}{\begin{vmatrix} 1-b+m & -m' \\ -m & 1-b'+m' \end{vmatrix}} \quad Y' = \frac{\begin{vmatrix} 1-b+m & a + I_0 + M'_0 - M_0 \\ -m & a' + I'_0 + M_0 - M'_0 \end{vmatrix}}{\begin{vmatrix} 1-b+m & -m' \\ -m & 1-b'+m' \end{vmatrix}}.$$

4.7 Input-output table

Assumption: Technologies are all fixed proportional, that is, to produce one unit of product X_i, you need a_{ji} units of X_j.

IO table: $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$.

Column i represents the coefficients of inputs needed to produce one unit of X_i.

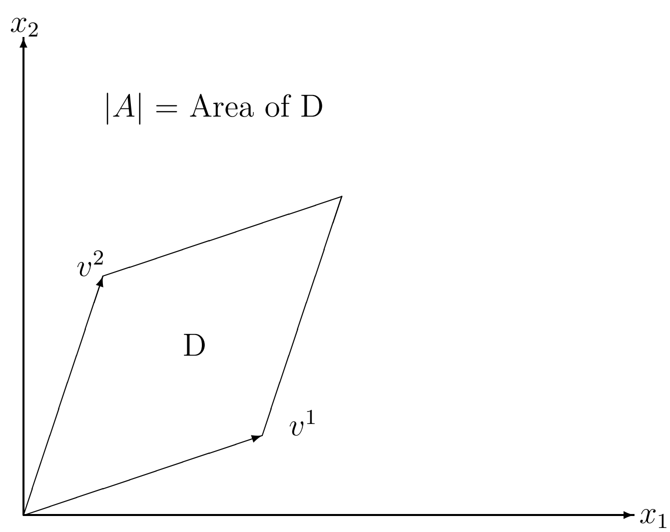
Suppose we want to produce a list of outputs $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, we will need a list of inputs

$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_2 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{pmatrix}$. The net output is $x - Ax = (I - A)x$.

If we want to produce a net amount of $d = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$, then since $d = (I - A)x$, $x = (I - A)^{-1}d$.

4.8 A geometric interpretation of determinants

Because of properties 2 and 4, the determinant function $D(v^1, \ldots, v^n)$ is called an alternative linear n-form of R^n. It is equal to the volume of the parallelepiped formed by the vectors $\{v^1, \ldots, v^n\}$. For $n = 2$, $|A|$ is the area of the parallelogram formed by $\left\{ \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix}, \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} \right\}$. See the diagram:



If the determinant is 0, then the volume is 0 and the vectors are linearly dependent, one of them must be a linear combination of others. Therefore, an inverse mapping does not exist, A^{-1} does not exist, and A is singular.

4.9 Rank of a matrix and solutions of $Ax = d$ when $|A| = 0$

Rank(A) = the maximum # of independent vectors in $A = \{v^1, \ldots, v^n\}$ = dim(Range Space of T_A).
Rank(A) = the size of the largest non-singular square submatrices of A.
Examples: $\text{Rank}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 2$. $\text{Rank}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 2$ because $\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$ is non-singular.
Property 1: $\text{Rank}(AB) \le \min\{\text{Rank}(A), \text{Rank}(B)\}$.
Property 2: dim(Null Space of T_A) + dim(Range Space of T_A) = n.

Consider the simultaneous equation $Ax = d$. When $|A| = 0$, there exists a row of A that is a linear combination of other rows

and Rank(A) $< n$. First, form the augmented matrix $M \equiv [A \vdots d]$ and calculate the rank of M. There are two cases.

Case 1: Rank(M) = Rank (A).
In this case, some equations are linear combinations of others (the equations are dependent) and can be removed without changing the solution space. There will be more variables than equations after removing these equations. Hence, there will be infinite number of solutions.
Example: $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$. $\text{Rank}(A) = \text{Rank}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1 = \text{Rank}(M) = \text{Rank}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$.
The second equation is just twice the first equation and can be discarded. The solutions are $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3-2k \\ k \end{pmatrix}$ for any k. On x_1-x_2 space, the two equations are represented by the same line and every point on the line is a solution.

Case 2: Rank(M) = Rank(A) + 1.
In this case, there exists an equation whose LHS is a linear combination of the LHS of other equations but whose RHS is different from the same linear combination of the RHS of other equations. Therefore, the equation system is contradictory and there will be no solutions.
Example: $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$. $\text{Rank}(A) = \text{Rank}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1 < \text{Rank}(M) = \text{Rank}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} = 2$.
Multiplying the first equation by 2, $2x_1 + 4x_2 = 6$, whereas the second equation says $2x_1 + 4x_2 = 7$. Therefore, it is impossible to have any $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ satisfying both equations simultaneously. On x_1-x_2 space, the two equations are represented by two parallel lines and cannot have any intersection points.

4.10 Problems

1. Suppose $v_1 = (1,2,3)'$, $v_2 = (2,3,4)'$, and $v_3 = (3,4,5)'$. Is $\{v_1, v_2, v_3\}$ linearly independent? Why or why not?

2.. Find the inverse of A $= \begin{bmatrix} 6 & 5 \\ 8 & 7 \end{bmatrix}$.

3. Given the 3×3 matrix A $= \begin{bmatrix} 2 & 1 & 6 \\ 5 & 3 & 4 \\ 8 & 9 & 7 \end{bmatrix}$,

(a) calculate the cofactors C_{11}, C_{21}, C_{31},

(b) use Laplace expansion theorem to find |A|,

(c) and use Cramer's rule to find X_1 of the following equation system:

$$\begin{bmatrix} 2 & 1 & 6 \\ 5 & 3 & 4 \\ 8 & 9 & 7 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

(Hint: Make use of the results of (a).)

4. Use Cramer's rule to solve the national-income model

$$C = a + b(Y - T) \quad (1)$$

$$T = -t_0 + t_1 Y \quad (2)$$

$$Y = C + I_0 + G \quad (3)$$

5. Let $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(a) Find AA and AAA.

(b) Let $x = (1,2,3)'$, compute Ax, AAx, and $AAAx$.

(c) Find $Rank[A]$, $Rank[AA]$, and $Rank[AAA]$.

6. Let $X = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$.

(a) Find $X'X$ and $(X'X)^{-1}$.

(b) Compute $X(X'X)^{-1}X'$ and $I - X(X'X)^{-1}X'$.

(c) Find $Rank[X(X'X)^{-1}X']$ and $Rank[I - X(X'X)^{-1}X']$.

7. A $= \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, B $= \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 1 \end{bmatrix}$, and C $= \begin{bmatrix} 1 & 2 & 4 \\ 3 & 6 & 12 \end{bmatrix}$.

(a) Find the ranks of A, B, and C.

(b) Use the results of (a) to determine whether the following system has any solution:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(c) Do the same for the following system:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}.$$

8. Let $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$, I the 2×2 identity matrix, and λ a scalar number.
(a) Find $|A - \lambda I|$. (Hint: It is a quadratic function of λ.)
(b) Determine Rank$(A - I)$ and Rank$(A - 4I)$. (Remark: $\lambda = 1$ and $\lambda = 4$ are the eigenvalues of A, that is, they are the roots of the equation $|A - \lambda I| = 0$, called the characteristic equation of A.)
(c) Solve the simultaneous equation system $(A - I)x = 0$ assuming that $x_1 = 1$. (Remark: The solution is called an eigenvector of A associated with the eigenvalue $\lambda = 1$.)
(d) Solve the simultaneous equation system $(A - 4I)y = 0$ assuming that $y_1 = 1$.
(e) Determine whether the solutions x and y are linearly independent.

5 Differential Calculus and Comparative Statics

As seen in the last chapter, a linear economic model can be represented by a matrix equation $Ax = d(y)$ and solved using Cramer's rule, $x = A^{-1}d(y)$. On the other hand, a closed form solution $x = x(y)$ for a nonlinear economic model is, in most applications, impossible to obtain. For general nonlinear economic models, we use differential calculus (implicit function theorem) to obtain the derivatives of endogenous variables with respect to exogenous variables $\dfrac{\partial x_i}{\partial y_j}$:

$$\begin{aligned} f_1(x_1, \ldots, x_n; y_1, \ldots, y_m) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n; y_1, \ldots, y_m) &= 0 \end{aligned}$$

$$\Rightarrow \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_m} \end{pmatrix} = - \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{pmatrix}.$$

$$\Rightarrow \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_m} \end{pmatrix} = - \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{pmatrix}.$$

Each $\dfrac{\partial x_i}{\partial y_j}$ represents a cause-effect relationship. If $\dfrac{\partial x_i}{\partial y_j} > 0$ (< 0), then x_i will increase (decrease) when y_j increases. Therefore, instead of computing $x_i = x_i(y)$, we want to determine the sign of $\dfrac{\partial x_i}{\partial y_j}$ for each i-j pair. In the following, we will explain how it works.

5.1 Differential Calculus

$$x = f(y) \Rightarrow f'(y^*) = \left.\frac{dx}{dy}\right|_{y=y^*} \equiv \lim_{\Delta y \to 0} \frac{f(y^* + \Delta y) - f(y^*)}{\Delta y}.$$

On y-x space, $x = f(y)$ is represented by a curve and $f'(y^*)$ represents the slope of the tangent line of the curve at the point $(y, x) = (y^*, f(y^*))$.
Basic rules:

1. $x = f(y) = k$, $\dfrac{dx}{dy} = f'(y) = 0$.
2. $x = f(y) = y^n$, $\dfrac{dx}{dy} = f'(y) = ny^{n-1}$.
3. $x = cf(y)$, $\dfrac{dx}{dy} = cf'(y)$.

4. $x = f(y) + g(y)$, $\dfrac{dx}{dy} = f'(y) + g'(y)$.

5. $x = f(y)g(y)$, $\dfrac{dx}{dy} = f'(y)g(y) + f(y)g'(y)$.

6. $x = f(y)/g(y)$, $\dfrac{dx}{dy} = \dfrac{f'(y)g(y) - f(y)g'(y)}{(g(y))^2}$.

7. $x = e^{ay}$, $\dfrac{dx}{dy} = ae^{ay}$. $\quad x = \ln y$, $\dfrac{dx}{dy} = \dfrac{1}{y}$.

8. $x = \sin y$, $\dfrac{dx}{dy} = \cos y$. $\quad x = \cos y$, $\dfrac{dx}{dy} = -\sin y$.

Higher order derivatives:

$f''(y) \equiv \dfrac{d}{dy}\left(\dfrac{d}{dy} f(y)\right) = \dfrac{d^2}{dy^2} f(y)$, $f'''(y) \equiv \dfrac{d}{dy}\left(\dfrac{d^2}{dy^2} f(y)\right) = \dfrac{d^3}{dy^3} f(y)$.

5.2 Partial derivatives

In many cases, x is a function of several y's: $x = f(y_1, y_2, \ldots, y_n)$. The partial derivative of x with respect to y_i evaluated at $(y_1, y_2, \ldots, y_n) = (y_1^*, y_2^*, \ldots, y_n^*)$ is

$$\left.\frac{\partial x}{\partial y_i}\right|_{(y_1^*, y_2^*, \ldots, y_n^*)} \equiv \lim_{\Delta y_i \to 0} \frac{f(y_1^*, \ldots, y_i^* + \Delta y_i, \ldots, y_n^*) - f(y_1^*, \ldots, y_i^*, \ldots, y_n^*)}{\Delta y_i},$$

that is, we regard all other independent variables as constant (f as a function of y_i only) and take derivative.

9. $\dfrac{\partial x_1^n x_2^m}{\partial x_1} = n x_1^{n-1} x_2^m$.

Higher order derivatives: We can define higher order derivatives as before. For the case with two independent variables, there are 4 second order derivatives:

$$\frac{\partial}{\partial y_1}\frac{\partial x}{\partial y_1} = \frac{\partial^2 x}{\partial y_1^2}, \quad \frac{\partial}{\partial y_2}\frac{\partial x}{\partial y_1} = \frac{\partial^2 x}{\partial y_2 \partial y_1}, \quad \frac{\partial}{\partial y_1}\frac{\partial x}{\partial y_2} = \frac{\partial^2 x}{\partial y_1 \partial y_2}, \quad \frac{\partial}{\partial y_2}\frac{\partial x}{\partial y_2} = \frac{\partial^2 x}{\partial y_2^2}.$$

Notations: f_1, f_2, f_{11}, f_{12}, f_{21}, f_{22}.

$\nabla f \equiv \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$: Gradient vector of f.

$H(f) \equiv \begin{pmatrix} f_{11} & \ldots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{n1} & \ldots & f_{nn} \end{pmatrix}$: second order derivative matrix, called Hessian of f.

Equality of cross-derivatives: If f is twice continously differentiable, then $f_{ij} = f_{ji}$ and $H(f)$ is symmetric.

5.3 Economic concepts similar to derivatives

Elasticity of X_i w.r.t. Y_j: $E_{X_i, Y_j} \equiv \dfrac{Y_j}{X_i}\dfrac{\partial X_i}{\partial Y_j}$, the percentage change of X_i when Y_j increases by 1 %. Example: $Q_d = D(P)$, $E_{Q_d, P} = \dfrac{P}{Q_d}\dfrac{dQ_d}{dP}$

Basic rules: 1. $E_{X_1X_2,Y} = E_{X_1,Y} + E_{X_2,Y}$, 2. $E_{X_1/X_2,Y} = E_{X_1,Y} - E_{X_2,Y}$, 3. $E_{Y,X} = 1/E_{X,Y}$.

Growth rate of $X = X(t)$: $G_X \equiv \frac{1}{X}\frac{dX}{dt}$, the percentage change of X per unit of time.

5.4 Mean value and Taylor's Theorems

Continuity theorem: If $f(y)$ is continuous on the interval $[a,b]$ and $f(a) \le 0$, $f(b) \ge 0$, then there exists a $c \in [a,b]$ such that $f(c) = 0$.
Rolle's theorem: If $f(y)$ is continuous on the interval $[a,b]$ and $f(a) = f(b) = 0$, then there exists a $c \in (a,b)$ such that $f'(c) = 0$.

Mean value theorem: If $f(y)$ is continously differentiable on $[a,b]$, then there exists a $c \in (a,b)$ such that

$$f(b) - f(a) = f'(c)(b-a) \quad \text{or} \quad \frac{f(b)-f(a)}{b-a} = f'(c).$$

Taylor's Theorem: If $f(y)$ is $k+1$ times continously differentiable on $[a,b]$, then for each $y \in [a,b]$, there exists a $c \in (a,y)$ such that

$$f(y) = f(a) + f'(a)(y-a) + \frac{f''(a)}{2!}(y-a)^2 + \ldots + \frac{f^{(k)}(a)}{k!}(y-a)^k + \frac{f^{(k+1)}(c)}{(k+1)!}(y-a)^{k+1}.$$

5.5 Concepts of differentials and applications

Let $x = f(y)$. Define $\Delta x \equiv f(y+\Delta y) - f(y)$, called the finite difference of x.
Finite quotient: $\frac{\Delta x}{\Delta y} = \frac{f(y+\Delta y) - f(y)}{\Delta y} \Rightarrow \Delta x = \frac{\Delta x}{\Delta y}\Delta y$.
dx, dy: Infinitesimal changes of x and y, $dx, dy > 0$ (so that we can divid something by dx or by dy) but $dx, dy < a$ for any positive real number a (so that $\Delta y \to dy$).
Differential of $x = f(y)$: $dx = df = f'(y)dy$.

Chain rule: $x = f(y),\ y = g(z) \Rightarrow x = f(g(z))$,
$dx = f'(y)dy,\ dy = g'(z)dz \Rightarrow dx = f'(y)g'(z)dz$. $\quad \frac{dx}{dz} = f'(y)g'(z) = f'(g(z))g'(z)$.
Example: $x = (z^2+1)^3 \Rightarrow x = y^3,\ y = z^2+1 \Rightarrow \frac{dx}{dz} = 3y^2 2z = 6z(z^2+1)^2$.
Inverse function rule: $x = f(y),\ \Rightarrow y = f^{-1}(x) \equiv g(x)$,
$dx = f'(y)dy,\ dy = g'(x)dx \Rightarrow dx = f'(y)g'(x)dx$. $\quad \frac{dy}{dx} = g'(x) = \frac{1}{f'(y)}$.
Example: $x = \ln y \Rightarrow y = e^x \Rightarrow \frac{dx}{dy} = \frac{1}{e^x} = \frac{1}{y}$.

5.6 Concepts of total differentials and applications

Let $x = f(y_1, y_2)$. Define $\Delta x \equiv f(y_1 + \Delta y_1, y_2 + \Delta y_2) - f(y_1, y_2)$, called the finite difference of x.

$$\begin{aligned}
\Delta x &= f(y_1 + \Delta y_1, y_2 + \Delta y_2) - f(y_1, y_2) \\
&= f(y_1 + \Delta y_1, y_2 + \Delta y_2) - f(y_1, y_2 + \Delta y_2) + f(y_1, y_2 + \Delta y_2) - f(y_1, y_2) \\
&= \frac{f(y_1 + \Delta y_1, y_2 + \Delta y_2) - f(y_1, y_2 + \Delta y_2)}{\Delta y_1}\Delta y_1 + \frac{f(y_1, y_2 + \Delta y_2) - f(y_1, y_2)}{\Delta y_2}\Delta y_2 \\
dx &= f_1(y_1, y_2)dy_1 + f_2(y_1, y_2)dy_2.
\end{aligned}$$

dx, $dy = \begin{pmatrix} dy_1 \\ \vdots \\ dy_n \end{pmatrix}$: Infinitesimal changes of x (endogenous), $y_1, \ldots, y_n$ (exogenous).

Total differential of $x = f(y_1, \ldots, y_n)$:

$$dx = df = f_1(y_1, \ldots, y_n)dy_1 + \ldots + f_n(y_1, \ldots, y_n)dy_n = (f_1, \ldots, f_n)\begin{pmatrix} dy_1 \\ \vdots \\ dy_n \end{pmatrix} = (\nabla f)'dy.$$

Implicit function rule:

In many cases, the relationship between two variables are defined implicitly. For example, the indifference curve $U(x_1, x_2) = \bar{U}$ defines a relationship between x_1 and x_2. To find the slope of the curve $\dfrac{dx_2}{dx_1}$, we use implicit function rule.

$$dU = U_1(x_1, x_2)dx_1 + U_2(x_1, x_2)dx_2 = d\bar{U} = 0 \Rightarrow; \frac{dx_2}{dx_1} = -\frac{U_1(x_1, x_2)}{U_2(x_1, x_2)}.$$

Example: $U(x_1, x_2) = 3x_1^{\frac{1}{3}} + 3x_2^{\frac{1}{3}} = 6$ defines an indifference curve passing through the point $(x_1, x_2) = (1, 1)$. The slope (Marginal Rate of Substitution) at $(1, 1)$ can be calculated using implicit function rule.

$$\frac{dx_2}{dx_1} = -\frac{U_1}{U_2} = -\frac{x_1^{-\frac{2}{3}}}{x_2^{-\frac{2}{3}}} = -\frac{1}{1} = -1.$$

Multivariate chain rule:

$$x = f(y_1, y_2),\ y_1 = g^1(z_1, z_2),\ y_2 = g^2(z_1, z_2),\ \Rightarrow x = f(g^1(z_1, z_2), g^2(z_1, z_2)) \equiv H(z_1, z_2).$$

We can use the total differentials dx, dy_1, dy_2 to find the derivative $\dfrac{\partial x}{\partial z_1}$.

$$dx = (f_1, f_2)\begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} = (f_1, f_2)\begin{pmatrix} g_1^1 & g_2^1 \\ g_1^2 & g_2^2 \end{pmatrix}\begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix} = (f_1 g_1^1 + f_2 g_1^2, f_1 g_2^1 + f_2 g_2^2)\begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}.$$

$$\Rightarrow \frac{\partial x}{\partial z_1} = \frac{\partial H}{\partial z_1} = f_1 g_1^1 + f_2 g_1^2, \quad \frac{\partial x}{\partial z_2} = \frac{\partial H}{\partial z_2} = f_1 g_2^1 + f_2 g_2^2.$$

Example: $x = y_1^6 y_2^7$, $y_1 = 2z_1 + 3z_2$, $y_2 = 4z_1 + 5z_2$, $\frac{\partial x}{\partial z_1} = 6y_1^5 y_2^7(2) + 7y_1^6 y_2^6(4)$.

Total derivative:

$$x = f(y_1, y_2),\ y_2 = h(y_1),\ x = f(y_1, h(y_1)) \equiv g(y_1),$$
$$\Rightarrow dx = f_1 dy_1 + f_2 dy_2 = f_1 dy_1 + f_2 h' dy_1 = (f_1 + f_2 h')dy_1.$$

Total derivative: $\left.\frac{dx}{dy_1}\right|_{y_2=h(y_1)} = f_1 + f_2 h'$.

Partial derivative (direct effect of y_1 on x): $\frac{\partial x}{\partial y_1} = \frac{\partial f}{\partial y_1} = f_1(y_1, y_2)$.

Indirect effect through y_2: $\frac{\partial x}{\partial y_2}\frac{dy_2}{dy_1} = f_2 h'$.

Example: Given the utility function $U(x_1, x_2) = 3x_1^{\frac{1}{3}} + 3x_2^{\frac{1}{3}}$, the MRS at a point (x_1, x_2) is $m(x_1, x_2) = \frac{dx_2}{dx_1} = -\frac{U_1(x_1, x_2)}{U_2(x_1, x_2)} = -\frac{x_1^{-\frac{2}{3}}}{x_2^{-\frac{2}{3}}}$. The rate of change of MRS w.r.t. x_1 along the indifference curve passing through $(1, 1)$ is a total derivative

$$\left.\frac{dm}{dx_1}\right|_{3x_1^{1/3}+3x_2^{1/3}=6} \left(= \left.\frac{d^2 x_2}{dx_1^2}\right|_{3x_1^{1/3}+3x_2^{1/3}=6} \right) = \frac{\partial m}{\partial x_1} + \frac{\partial m}{\partial x_2}\frac{dx_2}{dx_1} = \frac{\partial m}{\partial x_1} + \frac{\partial m}{\partial x_2}\left(-\frac{x_1^{-\frac{2}{3}}}{x_2^{-\frac{2}{3}}} \right).$$

5.7 Inverse function theorem

In Lecture 3, we discussed a linear mapping $x = Ay$ and its inverse mapping $y = A^{-1}x$ when $|A| \neq 0$.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{a_{22}}{|A|} & \frac{-a_{12}}{|A|} \\ \frac{-a_{21}}{|A|} & \frac{a_{11}}{|A|} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Therefore, for a linear mapping with $|A| \neq 0$, an 1-1 inverse mapping exists and the partial derivatives are given by the inverse matrix of A. For example, $\partial x_1/\partial y_1 = a_{11}$ where$\partial y_1/\partial x_1 = \frac{a_{22}}{|A|}$ etc. The idea can be generalized to nonlinear mappings.

A general nonlinear mapping from R^n to R^n, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rightarrow x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, is represented by a vector function

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f^1(y_1, \ldots, y_n) \\ \vdots \\ f^n(y_1, \ldots, y_n) \end{pmatrix} \equiv F(y).$$

$$\text{Jacobian matrix: } J_F(y) \equiv \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} = \begin{pmatrix} f_1^1 & \cdots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \cdots & f_n^n \end{pmatrix}.$$

Jacobian: $\dfrac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)} \equiv |J_F(y)|$.

Inverse function theorem: If $x^* = F(y^*)$ and $|J_F(y^*)| \neq 0$ ($J_F(y^*)$ is non-singular), then $F(y)$ is invertible nearby x^*,

that is, there exists a function $G(x) \equiv \begin{pmatrix} g^1(x_1, \ldots, x_n) \\ \vdots \\ g^n(x_1, \ldots, x_n) \end{pmatrix}$ such that $y = G(x)$ if $x = F(y)$. In that case, $J_G(x^*) = (J_F(y^*))^{-1}$.

Reasoning:

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = \begin{pmatrix} f_1^1 & \cdots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \cdots & f_n^n \end{pmatrix} \begin{pmatrix} dy_1 \\ \vdots \\ dy_n \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} dy_1 \\ \vdots \\ dy_n \end{pmatrix} = \begin{pmatrix} g_1^1 & \cdots & g_n^1 \\ \vdots & \ddots & \vdots \\ g_1^n & \cdots & g_n^n \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = \begin{pmatrix} f_1^1 & \cdots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \cdots & f_n^n \end{pmatrix}^{-1} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}$$

Example: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = F(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$. $J_F(r, \theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$.
$J = |J_F| = r(\cos^2\theta + \sin^2\theta) = r > 0$, $\Rightarrow r = \sqrt{x_1^2 + x_2^2}$, $\theta = \tan^{-1}\dfrac{x_2}{x_1}$ and $J_G = (J_F)^{-1}$. When $r = 0$, $J = 0$ and the mapping is degenerate, i.e., the whole set $\{r = 0, -\pi \leq \theta < \pi\}$ is mapped to the origin $(0, 0)$, just like the case in Lecture 3 when the Null space is a line.

Notice that $g_1^1 \neq 1/(f_1^1)$ in general.

5.8 Implicit function theorem and comparative statics

Linear model: If all the equations are linear, the model can be represtned in matrix form as

$$Ax + By = c \Leftrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \ddots \\ 0 \end{pmatrix}.$$

If $|A| \neq 0$, then the solution is given by $x = -A^{-1}(By + c)$. The derivative matrix $[\partial x_i/\partial y_j]_{ij} = A^{-1}B$. Using total differentials of the equations, we can derive a similar derivative matrix for general nonlinear cases.

We can regard the LHS of a nonlinear economic model as a mapping from R^{n+m} to R^n:

$$\begin{pmatrix} f_1(x_1,\ldots,x_n;y_1,\ldots,y_m)=0 \\ \vdots \\ f_n(x_1,\ldots,x_n;y_1,\ldots,y_m)=0 \end{pmatrix} \Leftrightarrow F(x;y)=0.$$

Jacobian matrix: $J_x \equiv \begin{pmatrix} f_1^1 & \ldots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \ldots & f_n^n \end{pmatrix}$.

Implicit function theorem: If $F(x^*;y^*)=0$ and $|J_x(x^*;y^*)| \neq 0$ ($J_x(x^*;y^*)$ is non-singular), then $F(x;y)=0$ is solvable nearby $(x^*;y^*)$, that is, there exists a function $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x(y) = \begin{pmatrix} x_1(y_1,\ldots,y_m) \\ \vdots \\ x_n(y_1,\ldots,y_m) \end{pmatrix}$ such that $x^* = x(y^*)$ and $F(x(y);y)=0$. In that case,

$$\begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_m} \end{pmatrix} = -\begin{pmatrix} f_1^1 & \ldots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \ldots & f_n^n \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f^1}{\partial y_1} & \cdots & \frac{\partial f^1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial y_1} & \cdots & \frac{\partial f^n}{\partial y_m} \end{pmatrix}.$$

Reasoning:

$$\begin{pmatrix} df^1 \\ \vdots \\ df^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} f_1^1 & \ldots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \ldots & f_n^n \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} + \begin{pmatrix} \frac{\partial f^1}{\partial y_1} & \cdots & \frac{\partial f^1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial y_1} & \cdots & \frac{\partial f^n}{\partial y_m} \end{pmatrix} \begin{pmatrix} dy_1 \\ \vdots \\ dy_m \end{pmatrix} = 0$$

$$\Rightarrow \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = -\begin{pmatrix} f_1^1 & \ldots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \ldots & f_n^n \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f^1}{\partial y_1} & \cdots & \frac{\partial f^1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial y_1} & \cdots & \frac{\partial f^n}{\partial y_m} \end{pmatrix} \begin{pmatrix} dy_1 \\ \vdots \\ dy_m \end{pmatrix}.$$

Example: $f^1 = x_1^2 x_2 - y = 0$, $f^2 = 2x_1 - x_2 - 1 = 0$, When $y=1$, $(x_1,x_2)=(1,1)$ is an equilibrium. To calculate $\frac{dx_1}{dy}$ and $\frac{dx_2}{dy}$ at the equilibrium we use the implicit function theorem:

$$\begin{aligned} \begin{pmatrix} \frac{dx_1}{dy} \\ \frac{dx_2}{dy} \end{pmatrix} &= -\begin{pmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f^1}{\partial y} \\ \frac{\partial f^2}{\partial y} \end{pmatrix} \\ &= -\begin{pmatrix} 2x_1x_2 & x_1^2 \\ 2 & -1 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/2 \end{pmatrix}. \end{aligned}$$

5.9 Problems

1. Given the demand function $Q_d = (100/P) - 10$, find the demand elasticity η.

2. Given $Y = X_1^2 X_2 + 2X_1 X_2^2$, find $\partial Y/\partial X_1$, $\partial^2 Y/\partial X_1^2$, and $\partial^2 Y/\partial X_1 \partial X_2$, and the total differential DY.

3. Given $Y = F(X_1, X_2) + f(X_1) + g(X_2)$, find $\partial Y/\partial X_1$, $\partial^2 Y/\partial X_1^2$, and $\partial^2 Y/\partial X_1 \partial X_2$.

4. Given the consumption function $C = C(Y - T(Y))$, find dC/dY.

5. Given that $Q = D(q * e/P)$, find dQ/dP.

6. $Y = X_1^2 X_2$, $Z = Y^2 + 2Y - 2$, use chain rule to derive $\partial Z/\partial X_1$ and $\partial Z/\partial X_2$.

7. $Y_1 = X_1 + 2X_2$, $Y_2 = 2X_1 + X_2$, and $Z = Y_1 Y_2$, use chain rule to derive $\partial Z/\partial X_1$ and $\partial Z/\partial X_2$.

8. Let $U(X_1, X_2) = X_1 X_2^2 + X_1^2 X_2$ and $X_2 = 2X_1 + 1$, find the partial derivative $\partial U/\partial X_1$ and the total derivative dU/dX_1.

9. $X^2 + Y^3 = 1$, use implicit function rule to find dY/dX.

10. $X_1^2 + 2X_2^2 + Y^2 = 1$, use implicit function rule to derive $\partial Y/\partial X_1$ and $\partial Y/\partial X_2$.

11. $F(Y_1, Y_2, X) = Y_1 - Y_2 + X - 1 = 0$ and $G(Y_1, Y_2, X) = Y_1^2 + Y_2^2 + X^2 - 1 = 0$. use implicit function theorem to derive dY_1/dX and dY_2/dX.

12. In a Cournot quantity competition duopoly model with heterogeneous products, the demand functions are given by

$$Q_1 = a - P_1 - cP_2, \quad Q_2 = a - cP_1 - P_2; \qquad 1 \geq c > 0.$$

   (a) For what value of c can we invert the demand functions to obtain P_1 and P_2 as functions of Q_1 and Q_2?

   (b) Calculate the inverse demand functions $P_1 = P_1(Q_1, Q_2)$ and $P_2 = P_2(Q_1, Q_2)$.

   (c) Derive the total revenue functions $TR_1(Q_1, Q_2) = P_1(Q_1, Q_2)Q_1$ and $TR_2(Q_1, Q_2) = P_2(Q_1, Q_2)Q_2$.

13. In a 2-good market equilibrium model, the inverse demand functions are given by

$$P_1 = A_1 Q_1^{\alpha-1} Q_2^{\beta}, \quad P_2 = A_2 Q_1^{\alpha} Q_2^{\beta-1}; \qquad \alpha, \beta > 0.$$

   (a) Calculate the Jacobian matrix $\begin{pmatrix} \partial P_1/\partial Q_1 & \partial P_1/\partial Q_2 \\ \partial P_2/\partial Q_1 & \partial P_2/\partial Q_2 \end{pmatrix}$ and Jacobian $\dfrac{\partial(P_1, P_2)}{\partial(Q_1, Q_2)}$. What condition(s) should the parameters satisfy so that we can invert the functions to obtain the demand functions?

   (b) Derive the Jacobian matrix of the derivatives of (Q_1, Q_2) with respect to (P_1, P_2), $\begin{pmatrix} \partial Q_1/\partial P_1 & \partial Q_1/\partial P_2 \\ \partial Q_2/\partial P_1 & \partial Q_2/\partial P_2 \end{pmatrix}$.

5.10 Proofs of important theorems of differentiation

Rolle's theorem: If $f(x) \in C[a,b]$, $f'(x)$ exists for all $x \in (a,b)$, and $f(a) = f(b) = 0$, then there exists a $c \in (a,b)$ such that $f'(c) = 0$.
Proof:
Case 1: $f(x) \equiv 0 \; \forall x \in [a,b] \Rightarrow f'(x) = 0 \; \forall x \in (a,b)./$
Case 2: $f(x) \not\equiv 0 \quad \in [a,b] \Rightarrow \exists e, c$ such that $f(e) = m \leq f(x) \leq M = f(c)$ and $M > m$. Assume that $M \neq 0$ (otherwise $m \neq 0$ and the proof is similar). It is easy to see that $f'_-(c) \geq 0$ and $f'_+(c) \leq 0$. Therefore, $f'(c) = 0$. *Q.E.D.*

Mean Value theorem: If $f(x) \in C[a,b]$ and $f'(x)$ exists for all $x \in (a,b)$. Then there exists a $c \in (a,b)$ such that

$$f(b) - f(a) = f'(c)(b-a).$$

Proof:
Consider the function

$$\phi(x) \equiv f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a) \right].$$

It is clear that $\phi(x) \in C[a,b]$ and $\phi'(x)$ exists for all $x \in (a,b)$. Also, $\phi(a) = \phi(b) = 0$ so that the conditions of Rolle's Theorem are satisfied for $\phi(x)$. Hence, there exists a $c \in (a,b)$ such that $\phi'(c) = 0$, or

$$\phi'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0 \quad \Rightarrow f'(c) = \frac{f(b)-f(a)}{b-a} = 0.$$

Q.E.D.

Taylor's Theorem: If $f(x) \in C^r[a,b]$ and $f^{(r+1)}(x)$ exists for all $x \in (a,b)$. Then there exists a $c \in (a,b)$ such that

$$f(b) = f(a)+f'(a)(b-a)+\frac{1}{2}f''(a)(b-a)^2+\ldots+\frac{1}{r!}f^{(r)}(a)(b-a)^r+\frac{1}{(r+1)!}f^{(r+1)}(c)(b-a)^{r+1}.$$

Proof:
Define $\xi \in R$

$$\frac{(b-a)^{r+1}}{(r+1)!}\xi \equiv f(b) - \left[f(a) + f'(a)(b-a) + \frac{1}{2}f''(a)(b-a)^2 + \ldots + \frac{1}{r!}f^{(r)}(a)(b-a)^r \right].$$

Consider the function

$$\phi(x) \equiv f(b) - \left[f(x) + f'(x)(b-x) + \frac{1}{2}f''(x)(b-x)^2 + \ldots + \right.$$

$$\left. \frac{1}{r!}f^{(r)}(x)(b-x)^r + \frac{\xi}{(r+1)!}(b-x)^{r+1} \right].$$

It is clear that $\phi(x) \in C[a,b]$ and $\phi'(x)$ exists for all $x \in (a,b)$. Also, $\phi(a) = \phi(b) = 0$ so that the conditions of Rolle's Theorem are satisfied for $\phi(x)$. Hence, there exists a $c \in (a,b)$ such that $\phi'(c) = 0$, or

$$\phi'(c) = \frac{\xi - f^{(r+1)}(c)}{r!} = 0 \quad \Rightarrow f^{(r+1)}(c) = \xi.$$

Q.E.D.

Inverse Function Theorem: Let $E \subseteq R^n$ be an open set. Suppose $f: E \to R^n$ is $C^1(E)$, $a \in E$, $f(a) = b$, and $A = J(f(a))$, $|A| \neq 0$. Then there exist open sets $U, V \subset R^n$ such that $a \in U$, $b \in V$, f is one to one on U, $f(U) = V$, and $f^{-1}: V \to U$ is $C^1(U)$.

Proof:

(1. Find U.) Choose $\lambda \equiv |A|/2$. Since $f \in C^1(E)$, there exists a neighborhood $U \subseteq E$ with $a \in U$ such that $\|J(f(x)) - A\| < \lambda$.

(2. Show that $f(x)$ is one to one in U.) For each $y \in R^n$ define ϕ_y on E by $\phi_y(x) \equiv x + A^{-1}(y - f(x))$. Notice that $f(x) = y$ if and only if x is a fixed point of ϕ_y. Since $J(\phi_y(x)) = I - A^{-1}J(f(x)) = A^{-1}[A - J(f(x))] \Rightarrow \|J(\phi_y(x))\| < \dfrac{1}{2}$ on U. Therefore $\phi_y(x)$ is a contraction mapping and there exists at most one fixed point in U. Therefore, f is one to one in U.

(3. $V = f(U)$ is open so that f^{-1} is continuous.) Let $V = f(U)$ and $y_0 = f(x_0) \in V$ for $x_0 \in U$. Choose an open ball B about x_0 with radius ρ such that the closure $[B] \subseteq U$. To prove that V is open, it is enough to show that $y \in V$ whenever $\|y - y_0\| < \lambda\rho$. So fix y such that $\|y - y_0\| < \lambda\rho$. With ϕ_y defined above,

$$\|\phi_y(x_0) - x_0\| = \|A^{-1}(y - y_0)\| < \|A^{-1}\|\lambda\rho = \frac{\rho}{2}.$$

If $x \in [B] \subseteq U$, then

$$\|\phi_y(x) - x_0\| \le \|\phi_y(x) - \phi_y(x_0)\| + \|\phi_y(x_0) - x_0\| < \frac{1}{2}\|x - x_0\| + \frac{\rho}{2} \le \rho.$$

That is, $\phi_y(x) \in [B]$. Thus, $\phi_y(x)$ is a contraction of the complete space $[B]$ into itself. Hence, $\phi_y(x)$ has a unique fixed point $x \in [B]$ and $y = f(x) \in f([B]) \subset f(U) = V$.

(4. $f^{-1} \in C^{-1}$.) Choose $y_1, y_2 \in V$, there exist $x_1, x_2 \in U$ such that $f(x_1) = y_1$, $f(x_2) = y_2$.

$$\phi_y(x_2) - \phi_y(x_1) = x_2 - x_1 + A^{-1}(f(x_1) - f(x_2)) = (x_2 - x_1) - A^{-1}(y_2 - y_1).$$

$$\Rightarrow \|(x_2 - x_1) - A^{-1}(y_2 - y_1)\| \le \frac{1}{2}\|x_2 - x_1\| \Rightarrow \frac{1}{2}\|x_2 - x_1\| \le \|A^{-1}(y_2 - y_1)\| \le \frac{1}{2\lambda}\|y_2 - y_1\|$$

or $\|x_2 - x_1\| \le \dfrac{1}{\lambda}\|y_2 - y_1\|$. It follows that $(f')^{-1}$ exists locally about a. Since

$$f^{-1}(y_2) - f^{-1}(y_1) - (f')^{-1}(y_1)(y_2 - y_1) = (x_2 - x_1) - (f')^{-1}(y_1)(y_2 - y_1)$$
$$= -(f')^{-1}(y_1)[-f'(x_1)(x_2 - x_1) + f(x_2) - f(x_1)],$$

We have

$$\frac{\|f^{-1}(y_2) - f^{-1}(y_1) - (f')^{-1}(y_1)(y_2 - y_1)\|}{\|y_2 - y_1\|} \leq \frac{\|(f')^{-1}\|}{\lambda} \frac{\|f(x_2) - f(x_1) - f'(x_1)(x_2 - x_1)\|}{\|x_2 - x_1\|}.$$

As $y_2 \rightarrow y_1$, $x_2 \rightarrow x_1$. Hence $(f^{1-})'(y) = \{f'[f^{-1}(y)]\}$ for $y \in V$. Since f^{-1} is differentiable, it is continuous. Also, f' is continuous and its inversion, where it exists, is continuous. Therefore $(f^{-1})'$ is continuous or $f^{-1} \in C^1(V)$. *Q.E.D.*

Implicit Function Theorem: Let $E \subseteq R^{(n+m)}$ be an open set and $a \in R^n$, $b \in R^m$, $(a, b) \in E$. Suppose $f : E \rightarrow R^n$ is $C^1(E)$ and $f(a, b) = 0$, and $J(f(a, b)) \neq 0$. Then there exist open sets $A \subset R^n$ and $B \subset R^m$ $a \in A$ and $b \in B$, such that for each $x \in B$, there exists a unique $g(x) \in A$ such that $f(g(x), x) = 0$ and $g : B \rightarrow A$ is $C^1(B)$.

Proof:

Defining $F : R^{n+m} \rightarrow R^{n+m}$ by $F(x, y) \equiv (x, f(x, y))$. Note that since

$$J(F(a,b)) = \begin{pmatrix} \left(\frac{\partial x_i}{\partial x_j}\right)_{1 \leq i,j \leq n} & \left(\frac{\partial x_i}{\partial x_{n+j}}\right)_{1 \leq i \leq n, 1 \leq j \leq m} \\ \left(\frac{\partial f_i}{\partial x_j}\right)_{1 \leq i \leq m, 1 \leq j \leq n} & \left(\frac{\partial f_i}{\partial x_{n+j}}\right)_{1 \leq i,j \leq m} \end{pmatrix} = \begin{pmatrix} I & O \\ N & M \end{pmatrix},$$

$|J(F(a, b))| = |M| \neq 0$. By the Inverse Function Theorem there exists an open set $V \subseteq R^{n+m}$ containing $F(a, b) = (a, 0)$ and an open set of the form $A \times B \subseteq E$ containing (a, b), such that $F : A \times B \rightarrow V$ has a C^1 inverse $F^{-1} : V \rightarrow A \times B$. F^{-1} is of the form $F^{-1}(x, y) = (x, \phi(x, y))$ for some C^1 function ϕ. Define the projection $\pi : R^{n+m} \rightarrow R^m$ by $\pi(x, y) = y$. Then $\pi \circ F(x, y) = f(x, y)$. Therefore

$$f(x, \phi(x, y)) = f \circ F^{-1}(x, y) = (\pi \circ F) \circ F^{-1}(x, y) = \pi \circ (F \circ F^{-1})(x, y) = \pi(x, y) = y$$

and $f(x, \phi(x, 0)) = 0$. So, define $g : A \rightarrow B$ by $g(x) = \phi(x, 0)$. *Q.E.D.*

6 Comparative Statics – Economic applications

6.1 Partial equilibrium model

$$\begin{array}{lll} Q = D(P,Y) & \dfrac{\partial D}{\partial P} < 0,\ \dfrac{\partial D}{\partial Y} > 0 & \text{end. var: } Q,\ P. \\ Q = S(P) & S'(P) > 0 & \text{ex. var: } Y.. \end{array}$$

$$\begin{array}{ll} f^1(P,Q;Y) = Q - D(P,Y) = 0 & \dfrac{df^1}{dY} = \dfrac{dQ}{dY} - \dfrac{\partial D}{\partial P}\dfrac{dP}{dY} - \dfrac{\partial D}{\partial Y} = 0 \\ f^2(P,Q;Y) = Q - S(P) = 0 & \dfrac{df^2}{dY} = \dfrac{dQ}{dY} - S'(P)\dfrac{dP}{dY} = 0 \end{array}.$$

$$\begin{pmatrix} 1 & -\dfrac{\partial D}{\partial P} \\ 1 & -S'(P) \end{pmatrix} \begin{pmatrix} \dfrac{dQ}{dY} \\ \dfrac{dP}{dY} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial D}{\partial Y} \\ 0 \end{pmatrix}, \quad |J| = \begin{vmatrix} 1 & -\dfrac{\partial D}{\partial P} \\ 1 & -S'(P) \end{vmatrix} = -S'(P) + \dfrac{\partial D}{\partial P} < 0.$$

$$\frac{dQ}{dY} = \frac{\begin{vmatrix} \dfrac{\partial D}{\partial Y} & -\dfrac{\partial D}{\partial P} \\ 0 & -S'(P) \end{vmatrix}}{|J|} = \frac{-\dfrac{\partial D}{\partial Y}S'(P)}{|J|} > 0, \quad \frac{dP}{dY} = \frac{\begin{vmatrix} 1 & \dfrac{\partial D}{\partial Y} \\ 1 & 0 \end{vmatrix}}{|J|} = \frac{-\dfrac{\partial D}{\partial Y}}{|J|} > 0.$$

6.2 Income determination model

$$\begin{array}{lll} C = C(Y) & 0 < C'(Y) < 1. & \\ I = I(r) & I'(r) < 0 & \text{end. var. } C,\ Y,\ I \\ Y = C + I + \bar{G} & & \text{ex. var. } \bar{G},\ r. \end{array}$$

$$Y = C(Y) + I(r) + \bar{G} \Rightarrow dY = C'(Y)dY + I'(r)dr + d\bar{G} \Rightarrow dY = \frac{I'(r)dr + d\bar{G}}{1 - C'(Y)}.$$

$$\frac{\partial Y}{\partial r} = \frac{I'(r)}{1 - C'(Y)} < 0, \quad \frac{\partial Y}{\partial \bar{G}} = \frac{1}{1 - C'(Y)} > 0.$$

6.2.1 Income determination and trade

Consider an income determination model with import and export:

$$\begin{array}{rl} C = C(Y) \ \ 1 > C_y > 0, & I = \bar{I}, \\ M = M(Y,e) \ \ M_y > 0,\ \ M_e < 0 & X = X(Y^*, e), X_{y*} > 0,\ \ X_e > 0 \\ C + I + X - M = Y, & \end{array}$$

where import M is a function of domestic income and exchange rate e and export X is a function of exchange rate and foreign income Y^*, both are assumed here as exogenous variables. Substituting consumption, import, and export functions into the equilibrium condition, we have

$$C(Y) + \bar{I} + X(Y^*, e) - M(Y,e) = Y, \quad \Rightarrow F(Y,e,Y^*) \equiv C(Y) + \bar{I} + X(Y^*, e) - M(Y,e) - Y = 0..$$

Use implicit function rule to derive $\frac{\partial Y}{\partial \bar{I}}$ and determine its sign:

$$\frac{\partial Y}{\partial \bar{I}} = -\frac{F_I}{F_y} = \frac{1}{1 - C'(Y) + M_y}.$$

Use implicit function rule to derive $\frac{\partial Y}{\partial e}$ and determine its sign:

$$\frac{\partial Y}{\partial e} = -\frac{F_e}{F_y} = \frac{X_e - M_e}{1 - C'(Y) + M_y}.$$

Use implicit function rule to derive $\frac{\partial Y}{\partial Y^*}$ and determine its sign:

$$\frac{\partial Y}{\partial Y^*} = -\frac{F_{y^*}}{F_y} = \frac{X_{y^*}}{1 - C'(Y) + M_y}.$$

6.2.2 Interdependence of domestic and foreign income

Now extend the above income determination model to analyze the joint dependence of domestic income and foreign income:

$$C(Y) + \bar{I} + X(Y^*, e) - M(Y, e) = Y \qquad C^*(Y^*) + \bar{I}^* + X^*(Y, e) - M^*(Y^*, e) = Y^*,$$

with a similar assumption on the foreigner's consumption function: $1 > C^*_{y*} > 0$. Since domestic import is the same as foreigner's export and domestic export is foreigner's import, $X^*(Y, e) = M(Y, e)$ and $M^*(Y^*, e) = X(Y^*, e)$ and the system becomes:

$$C(Y) + \bar{I} + X(Y^*, e) - M(Y, e) = Y \qquad C^*(Y^*) + \bar{I}^* + M(Y, e) - X(Y^*, e) = Y^*,$$

Calculate the total differential of the system (Now Y^* becomes endogenous):

$$\begin{pmatrix} 1 - C' + M_y & -X_{y*} \\ -M_y & 1 - C^{*\prime} + X_{y*} \end{pmatrix} \begin{pmatrix} dY \\ dY^* \end{pmatrix} = \begin{pmatrix} (X_e - M_e)de + d\bar{I} \\ (M_e - X_e)de + d\bar{I}^* \end{pmatrix},$$

$$|J| = \begin{vmatrix} 1 - C' + M_y & -X_{y*} \\ -M_y & 1 - C^{*\prime} + X_{y*} \end{vmatrix} = (1 - C' + M_y)(1 - C^{*\prime} + M_{y*}) - M_y X_{y^*} > 0.$$

Use Cramer's rule to derive $\frac{\partial Y}{\partial e}$ and $\frac{\partial Y^*}{\partial e}$ and determine their signs:

$$dY = \frac{\begin{vmatrix} (X_e - M_e)de + d\bar{I} & -X_{y*} \\ (M_e - X_e)de + d\bar{I}^* & 1 - C^{*\prime} + X_{y*} \end{vmatrix}}{|J|}$$

$$= \frac{(X_e - M_e)(1 - C^{*\prime} + X_{y*} - X_{y*})de + (1 - C^{*\prime} + X_{y*})d\bar{I} + X_{y*}d\bar{I}^*}{|J|},$$

$$dY^* = \frac{\begin{vmatrix} 1-C'+M_y & (X_e-M_e)de+d\bar{I} \\ -M_y & (M_e-X_e)de+d\bar{I}^* \end{vmatrix}}{|J|}$$

$$= \frac{-(X_e-M_e)(1-C'+M_y-M_y)de+(1-C'+M_y)d\bar{I}^*+X_y d\bar{I}}{|J|}.$$

$$\frac{\partial Y}{\partial e} = \frac{(X_e-M_e)(1-C^{*\prime})}{|J|} > 0, \qquad \frac{\partial Y^*}{\partial e} = \frac{-(X_e-M_e)(1-C')}{|J|} < 0.$$

Derive $\frac{\partial Y}{\partial \bar{I}}$ and $\frac{\partial Y^*}{\partial \bar{I}}$ and determine their signs:

$$\frac{\partial Y}{\partial \bar{I}} = -\frac{1-C^{*\prime}+M_{y*}}{|J|} > 0, \quad \frac{\partial Y^*}{\partial \bar{I}} = \frac{M_y}{|J|} < 0.$$

6.3 IS-LM model

$$\begin{array}{lll} C = C(Y) & 0 < C'(Y) < 1 & M^d = L(Y,r) \quad \frac{\partial L}{\partial Y} > 0, \ \frac{\partial L}{\partial r} < 0 \\ I = I(r) & I'(r) < 0 & M^s = \bar{M} \\ Y = C + I + \bar{G} & & M^d = M^s. \end{array}$$

end. var: Y, C, I, r, M^d, M^s. ex. var: $\bar{G}$, $\bar{M}$.

$$\begin{array}{r} Y - C(Y) - I(r) = \bar{G} \\ L(Y,r) = \bar{M} \end{array} \quad \Rightarrow \quad \begin{array}{r} (1-C'(Y))dY - I'(r)dr = d\bar{G} \\ \frac{\partial L}{\partial Y}dY + \frac{\partial L}{\partial r}dr = d\bar{M} \end{array}$$

$$\begin{pmatrix} 1-C' & -I' \\ L_Y & L_r \end{pmatrix}\begin{pmatrix} dY \\ dr \end{pmatrix} = \begin{pmatrix} d\bar{G} \\ d\bar{M} \end{pmatrix}, \quad |J| = \begin{vmatrix} 1-C' & -I' \\ L_Y & L_r \end{vmatrix} = (1-C')L_r + I'L_Y < 0.$$

$$dY = \frac{\begin{vmatrix} d\bar{G} & -I' \\ d\bar{M} & L_r \end{vmatrix}}{|J|} = \frac{L_r d\bar{G} + I'd\bar{M}}{|J|}, \ dr = \frac{\begin{vmatrix} 1-C' & d\bar{G} \\ L_Y & d\bar{M} \end{vmatrix}}{|J|} = \frac{-L_Y d\bar{G} + (1-C')d\bar{M}}{|J|}$$

$$\frac{\partial Y}{\partial \bar{G}} = \frac{L_r}{|J|} > 0, \ \ \frac{\partial Y}{\partial \bar{M}} = \frac{I'}{|J|} > 0, \ \ \frac{\partial r}{\partial \bar{G}} = -\frac{L_Y}{|J|} > 0, \ \ \frac{\partial r}{\partial \bar{M}} = \frac{(1-C')}{|J|} < 0.$$

6.4 Two-market general equilibrium model

$$\begin{array}{llll} Q_{1d} = D^1(P_1,P_2) & D^1_1 < 0, & Q_{2d} = D^2(P_1,P_2) & D^2_2 < 0, \ D^2_1 > 0. \\ Q_{1s} = \bar{S}_1 & & Q_{2s} = S_2(P_2) & S_2'(P_2) > 0 \\ Q_{1d} = Q_{1s} & & Q_{2d} = Q_{2s} & \end{array}$$

end. var: Q_1, Q_2, P_1, P_2. ex. var: $\bar{S}_1$.

$$\begin{array}{l} D^1(P_1,P_2) = \bar{S}_1 \\ D^2(P_1,P_2) - S_2(P_2) = 0 \end{array} \quad \Rightarrow \quad \begin{array}{l} D^1_1 dP_1 + D^1_2 dP_2 = d\bar{S}_1 \\ D^2_1 dP_1 + D^2_2 dP_2 - S_2' dP_2 = 0. \end{array}$$

$$\begin{pmatrix} D_1^1 & D_2^1 \\ D_1^2 & D_2^2 - S_2' \end{pmatrix} \begin{pmatrix} dP_1 \\ dP_2 \end{pmatrix} = \begin{pmatrix} d\bar{S}_1 \\ 0 \end{pmatrix}, \quad |J| = \begin{vmatrix} D_1^1 & D_2^1 \\ D_1^2 & D_2^2 - S_2' \end{vmatrix} = D_1^1(D_2^2 - S_2') - D_2^1 D_1^2.$$

Assumption: $|D_1^1| > |D_2^1|$ and $|D_2^2| > |D_1^2|$ (own-effects dominate), $\Rightarrow |J| > 0$.

$$\frac{dP_1}{d\bar{S}_1} = \frac{D_2^2 - S_2'}{|J|} < 0, \quad \frac{dP_2}{d\bar{S}_1} = \frac{-D_1^2}{|J|} < 0$$

From $Q_{1s} = \bar{S}_1$, $\dfrac{\partial Q_1}{\partial \bar{S}_1} = 1$. To calculate $\dfrac{\partial Q_2}{\partial \bar{S}_1}$, we have to use chain rule:

$$\frac{\partial Q_2}{\partial \bar{S}_1} = S_2' \frac{dP_2}{d\bar{S}_1} < 0.$$

6.4.1 Car market

Suppose we want to analyze the effect of the price of used cars on the market of new cars. The demand for new cars is given by $Q_n = D^n(P_n; P_u)$, $\partial D^n/\partial P_n < 0$ $\partial D^n/\partial P_u > 0$, where Q_n is the quantity of new cars and P_n (P_u) the price of a new (used) car. The supply function of new cars is $Q_n = S(P_n)$, $S'(P_n) > 0$.

end. var: P_n, Q_n. ex. var: P_u.

$$D^n(P_n; P_u) = S(P_n); \quad \Rightarrow \frac{\partial D^n}{\partial P_n} dP_n + \frac{\partial D^n}{\partial P_u} dP_u = S'(P_n) dP_n.$$

$$\frac{dP_n}{dP_u} = \frac{\partial D^n/\partial P_u}{S'(P_n) - \partial D^n/\partial P_n} > 0, \quad \frac{dQ_n}{dP_u} = S'(P_n)\frac{dP_n}{dP_u} = \frac{S'(P_n)\partial D^n/\partial P_u}{S'(P_n) - \partial D^n/\partial P_n} > 0.$$

The markets for used cars and for new cars are actually interrelated. The demand for used cars is $Q_u = D^u(P_u, P_n)$, $\partial D^n/\partial P_u < 0$, $\partial D^n/\partial P_n > 0$. In each period, the quantity of used cars supplied is fixed, denoted by $\bar{Q}_u$. Instead of analyzing the effects of a change in P_u on the new car market, we want to know how a change in $\bar{Q}_u$ affects both markets.

end. var: P_n, Q_n, P_u, Q_u. ex. var: $\bar{Q}_u$.

$$Q_n = D^n(P_n, P_u), \quad Q_n = S(P_n); \qquad Q_u = D^u(P_u, P_n), \quad Q_u = \bar{Q}_u$$

$$\Rightarrow D^n(P_n, P_u) = S(P_n), \quad D^u(P_u, P_n) = \bar{Q}_u; \qquad D_n^n dP_n + D_u^n dP_u = S' dP_n, \quad D_n^u dP_n + D_u^u dP_u = \bar{Q}$$

$$\begin{pmatrix} D_n^n - S' & D_u^n \\ D_n^u & D_u^u \end{pmatrix} \begin{pmatrix} dP_n \\ dP_u \end{pmatrix} = \begin{pmatrix} 0 \\ d\bar{Q}_u \end{pmatrix}, \quad |J| = \begin{vmatrix} D_n^n - S' & D_u^n \\ D_n^u & D_u^u \end{vmatrix} = (D_n^n - S')D_u^u - D_u^n D_n^u.$$

Assumption: $|D_n^n| > |D_u^n|$ and $|D_u^u| > |D_n^u|$ (own-effects dominate), $\Rightarrow |J| > 0$.

$$\frac{dP_n}{d\bar{Q}_u} = \frac{-D_u^n}{|J|} < 0, \quad \frac{dP_u}{d\bar{Q}_u} = \frac{D_n^n - S'}{|J|} < 0.$$

From $Q_u = \bar{Q}_u$, $\dfrac{\partial Q_u}{\partial \bar{Q}_u} = 1$. To calculate $\dfrac{\partial Q_n}{\partial \bar{Q}_u}$, we have to use chain rule: $\dfrac{\partial Q_n}{\partial \bar{Q}_u} = S' \dfrac{dP_n}{d\bar{Q}_n} < 0$.

6.5 Classic labor market model

$$L = h(w) \qquad h' > 0 \qquad \text{labor supply function}$$
$$w = \text{MPP}_L = \frac{\partial Q}{\partial L} = F_L(K, L) \quad F_{LK} > 0,\ F_{LL} < 0 \quad \text{labor demand function.}$$

endogenous variables: L, w. exogenous variable: K.

$$\begin{array}{ll} L - h(w) = 0 & \Rightarrow \quad dL - h'(w)dw = 0 \\ w - F_L = 0 & \qquad dw - F_{LL}dL - F_{LK}dK = 0 \end{array}.$$

$$\begin{pmatrix} 1 & -h'(w) \\ -F_{LL} & 1 \end{pmatrix}\begin{pmatrix} dL \\ dw \end{pmatrix} = \begin{pmatrix} 0 \\ F_{LK}dK \end{pmatrix}, \quad |J| = \begin{vmatrix} 1 & -h'(w) \\ -F_{LL} & 1 \end{vmatrix} = 1 - h'(w)F_{LL} > 0.$$

$$\frac{dL}{dK} = \frac{\begin{vmatrix} 0 & -h'(w) \\ F_{LK} & 1 \end{vmatrix}}{|J|} = \frac{h'F_{LK}}{|J|} > 0, \quad \frac{dw}{dK} = \frac{\begin{vmatrix} 1 & 0 \\ -F_{LL} & F_{LK} \end{vmatrix}}{|J|} = \frac{F_{LK}}{|J|} > 0.$$

6.6 Problem

1. Let the demand and supply functions for a commodity be

$$\begin{array}{rcll} Q &=& D(P) & D'(P) < 0 \\ Q &=& S(P, t) & \partial S/\partial P > 0, \quad \partial S/\partial t < 0, \end{array}$$

   where t is the tax rate on the commodity.
   (a) Derive the total differentail of each equation.
   (b) Use Cramer's rule to compute dQ/dt and dP/dt.
   (c) Determine the sign of dQ/dt and dP/dt.
   (d) Use the $Q - P$ diagram to explain your results.

2. Suppose consumption C depends on total wealth W, which is predetermind, as well as on income Y. The IS-LM model becomes

$$\begin{array}{llll} C = C(Y, W) & 0 < C_Y < 1 & C_W > 0 & M^S = M \\ I = I(r) & I'(r) < 0 & & Y = C + I \\ M^D = L(Y, r) & L_Y > 0 & L_r < 0 & M^S = M^D \end{array}$$

   (a) Which variables are endogenous? Which are exogenous?
   (b) Which equations are behavioral/institutional, which are equilibrium conditions?

   The model can be reduced to

$$Y - C(Y, W) - I(r) = 0, \qquad L(Y, r) = M.$$

   (c) Derive the total differential for each of the two equations.
   (d) Use Cramer's rule to derive the effects of an increase in W on Y and r, *ie.*, derive $\partial Y/\partial W$ and $\partial r/\partial W$.
   (e) Determine the signs of $\partial Y/\partial W$ and $\partial r/\partial W$.

3. Consider a 2-industry (e.g. manufacturing and agriculture) general equilibrium model. The demand for manufacturing product consists of two components: private demand D^1 and goverment demand G. The agricultural products have only private demand D^2. Both D^1 and D^2 depend only on their own prices. Because each industry requires in its production process outputs of the other, the supply of each commodity depends on the price of the other commodity as well as on its own price. Therefore, the model may be written as follows:

$$\begin{aligned}
Q_1^d &= D_1(P_1) + G & D_1'(P_1) &< 0, & & \\
Q_2^d &= D_2(P_2) & D_2'(P_2) &< 0, & & \\
Q_1^s &= S^1(P_1, P_2) & S_1^1 &> 0, & S_2^1 < 0,&\ \ S_1^1 > |S_2^1|, \\
Q_2^s &= S^2(P_1, P_2) & S_1^2 &< 0, & S_2^2 > 0,&\ \ S_2^2 > |S_1^2|, \\
Q_1^d &= Q_1^s, & & & & \\
Q_2^d &= Q_2^s. & & & &
\end{aligned}$$

(a) Which variables are endogenous? Which are exogenous?
(b) Which equations are behavioral? Which are definitional? Which are equilibrium conditions?

4. The model can be reduced to

$$S^1(P_1, P_2) = D_1(P_1) + G \qquad\qquad S^2(P_1, P_2) = D_2(P_2)$$

(c) Compute the total differential of each equation.
(d) Use Cramer's rule to derive $\partial P_1/\partial G$ and $\partial P_2/\partial G$.
(e) Determine the signs of $\partial P_1/\partial G$ and $\partial P_2/\partial G$.
(f) Compute $\partial Q_1/\partial G$ and $\partial Q_2/\partial G$. (Hint: Use chain rule.)
(g) Give an economic interpretation of the results.

5. The demand functions of a 2-commodity market model are:

$$Q_1^d = D^1(P_1, P_2) \qquad\qquad Q_2^d = D^2(P_1, P_2).$$

The supply of the first commodity is given exogenously, ie., $Q_1^s = S_1$. The supply of the second commodity depends on its own price, $Q_2^s = S_2(P_2)$. The equilibrium conditions are:

$$Q_1^d = Q_1^s, \qquad\qquad Q_2^d = Q_2^s.$$

(a) Which variables are endogenous? Which are exogenous?
(b) Which equations are behavioral? Which are definitional? Which are equilibrium conditions?

The model above can be reduced to :

$$D^1(P_1, P_2) - S_1 = 0 \qquad\qquad D^2(P_1, P_2) - S_2(P_2) = 0$$

Suppose that both commodities are not Giffen good (hence, $D_i^i < 0, i = 1, 2$), that each one is a gross substitute for the other (ie., $D_j^i > 0, i \neq j$), that $|D_i^i| > |D_j^i|$, and that $S_2'(P_2) > 0$.
(c) Calculate the total differential of each equation of the reduced model.
(d) Use Cramer's rule to derive $\partial P_1/\partial S_1$ and $\partial P_2/\partial S_1$.
(e) Determine the signs of $\partial P_1/\partial S_1$ and $\partial P_2/\partial S_1$.
(f) Compute $\partial Q_1/\partial S_1$ and $\partial Q_2/\partial S_1$ and determine their signs.

6. The demand functions for fish and chicken are as follows:

$$Q_F^d = D_F(P_F - P_C), D_F' < 0$$
$$Q_C^d = D_C(P_C - P_F), D_C' < 0$$

where P_F, P_C are price of fish and price of chicken respectively. The supply of fish depends on the number of fishermen (N) as well as its price P_F: $Q_F^s = F(P_F, N)$, $F_{P_F} > 0$, $F_N > 0$. The supply of chicken depends only on its price P_C: $Q_C^s = C(P_C)$, $C' > 0$. The model can be reduced to

$$\begin{aligned} D_F(P_F - P_C) &= F(P_F, N) \\ D_C(P_C - P_F) &= C(P_C) \end{aligned}$$

(a) Find the total differential of the reduced system.
(b) Use Cramer's rule to find dP_F/dN and dP_C/dN.
(c) Determine the signs of dP_F/dN and dP_C/dN. What is the economic meaning of your results?
(d) Find dQ_C/dN.

7. In a 2-good market equilibrium model, the inverse demand functions are given by

$$P_1 = U_1(Q_1, Q_2), \qquad P_2 = U_2(Q_1, Q_2);$$

where $U_1(Q_1, Q_2)$ and $U_2(Q_1, Q_2)$ are the partial derivatives of a utility function $U(Q_1, Q_2)$ with respect to Q_1 and Q_2, respectively.

(a) Calculate the Jacobian matrix $\begin{pmatrix} \partial P_1/\partial Q_1 & \partial P_1/\partial Q_2 \\ \partial P_2/\partial Q_1 & \partial P_2/\partial Q_2 \end{pmatrix}$ and Jacobian $\dfrac{\partial(P_1, P_2)}{\partial(Q_1, Q_2)}$. What condition(s) should the parameters satisfy so that we can invert the functions to obtain the demand functions?

(b) Derive the Jacobian matrix of the derivatives of (Q_1, Q_2) with respect to (P_1, P_2), $\begin{pmatrix} \partial Q_1/\partial P_1 & \partial Q_1/\partial P_2 \\ \partial Q_2/\partial P_1 & \partial Q_2/\partial P_2 \end{pmatrix}$.

(c) Suppose that the supply functions are

$$Q_1 = a^{-1}P_1, \qquad Q_2 = P_2,$$

and Q_1^* and Q_2^* are market equilibrium quantities. Find the comparative statics $\dfrac{dQ_1^*}{da}$ and $\dfrac{dQ_2^*}{da}$. (Hint: Eliminate P_1 and P_2.)

8. In a 2-good market equilibrium model with a sales tax of t dollars per unit on product 1, the model becomes

$$D^1(P_1 + t, P_2) = Q_1 = S_1(P_1), \quad D^2(P_1 + t, P_2) = Q_2 = S_2(P_2).$$

Suppose that $D_i^i < 0$ and $S_i' > 0$, $|D_i^i| > |D_j^i|$, $i \neq j$, $i, j = 1, 2$.

(a) Calculate dP_2/dt.

(b) Calculate dQ_2/dt.

(c) Suppose that $D_j^i > 0$. Determine the signs of dP_2/dt and dQ_2/dt.

(d) Suppose that $D_j^i < 0$. Determine the signs of dP_2/dt and dQ_2/dt.

(e) Explain your results in economics.

7 Optimization

A behavioral equation is a summary of the decisions of a group of economic agents in a model. A demand (supply) function summarizes the consumption (production) decisions of consumers (producers) under different market prices, etc. The derivative of a behavioral function represents how agents react when an independent variable changes. In the last chapter, when doing comparative static analysis, we always assumed that the signs of derivatives of a behavioral equation in a model are known. For example, $D'(P) < 0$ and $S'(P) > 0$ in the partial market equilibrium model, $C'(Y) > 0$, $I'(r) < 0$, $L_y > 0$, and $L_r < 0$ in the IS-LM model. In this chapter, we are going to provide a theoretical foundation for the determination of these signs.

7.1 Neoclassic methodology

Neoclassic assumption: An agent, when making decisions, has an objective function in mind (or has well defined preferences). The agent will choose a feasible decision such that the objective function is maximized.
A consumer will choose the quantity of each commodity within his/her budget constraints such that his/her utility function is maximized. A producer will choose to supply the quantity such that his profit is maximizaed.
Remarks: (1) Biological behavior is an alternative assumption, sometimes more appropriate, (2) Sometimes an agent is actually a group of people with different personalities like a company and we have to use game theoretic equilibrium concepts to characterize the collective behavior.

Maximization $\Rightarrow$ Behavioral equations
Game equilibrium $\Rightarrow$ Equilibrium conditions

$x_1, \ldots, x_n$: variables determined by the agent (endogenous variables).
$y_1, \ldots, y_m$: variables given to the agent (exogenous variables).
Objective function: $f(x_1, \ldots, x_n; y_1, \ldots, y_m)$.
Opportunity set: the agent can choose only $(x_1, \ldots, x_n)$ such that $(x_1, \ldots, x_n; y_1, \ldots, y_m) \in A \subset R^{n+m}$. A is usually defined by inequality constriants.

$$\max_{x_1,\ldots,x_n} f(x_1, \ldots, x_n; y_1, \ldots, y_m) \text{ subject to } \begin{cases} g^1(x_1, \ldots, x_n; y_1, \ldots, y_m) \geq 0 \\ \vdots \\ g^k(x_1, \ldots, x_n; y_1, \ldots, y_m) \geq 0. \end{cases}$$

Solution (behavioral equations): $x_i = x_i(y_1, \ldots, y_m)$, $i = 1, \ldots, n$ (derived from FOC).
$\partial x_i / \partial y_j$: derived by the comparative static method (sometimes its sign can be determined from SOC).

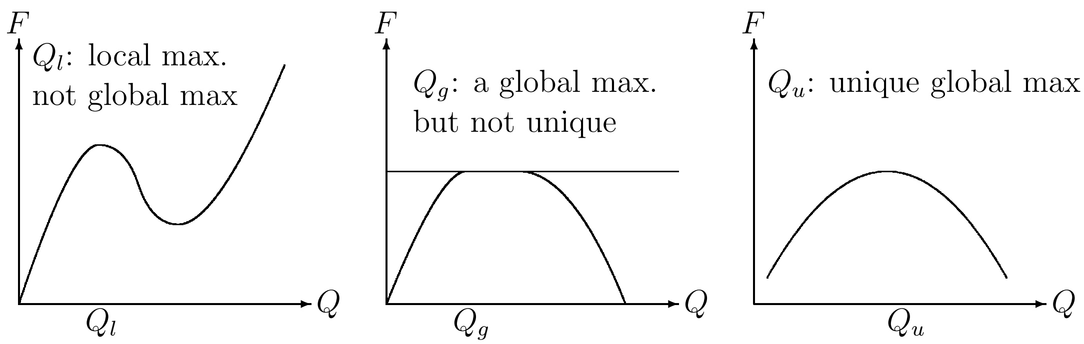
Example 1: A consumer maximizes his utility function $U(q_1, q_2)$ subject to the budget constraint $p_1 q_1 + p_2 q_2 = m \Rightarrow$ demand functions $q_1 = D^1(p_1, p_2, m)$ and $q_2 = D^2(p_1, p_2, m)$.

Example 2: A producer maximizes its profit $\Pi(Q;P) = PQ - C(Q)$ where $C(Q)$ is the cost of producing Q units of output $\Rightarrow$ the supply function $Q = S(P)$.

one endogenous variable: this chapter.
n endogenous variables without constraints: next
n endogenous variables with equality constraints:
Nonlinear programming: n endogenous variables with inequality constraints
Linear programming: Linear objective function with linear inequality constraints
Game theory: more than one agents with different objective functions

7.2 Different concepts of maximization

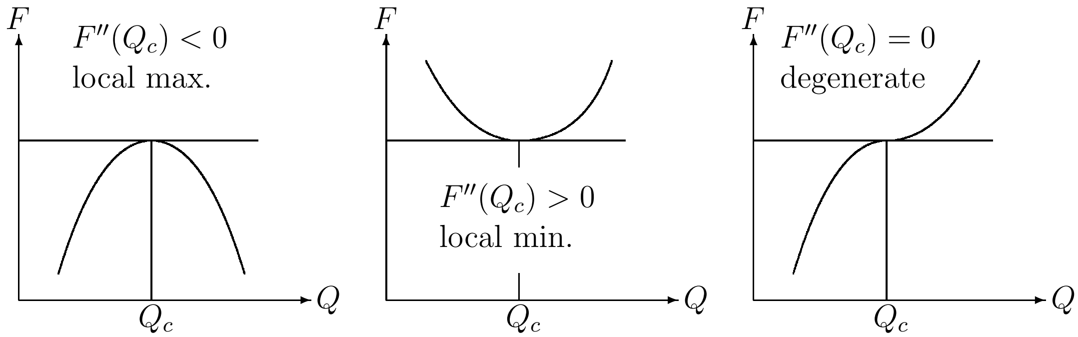
Suppose that a producer has to choose a Q to maximize its profit $\pi = F(Q)$: $\max_Q F(Q)$. Assume that $F'(Q)$ and $F''(Q)$ exist.
A **local** maximum Q_l: there exists $\epsilon > 0$ such that $F(Q_l) \geq F(Q)$ for all $Q \in (Q_l - \epsilon, Q_l + \epsilon)$.
A **global** maximum Q_g: $F(Q_g) \geq F(Q)$ for all Q.
A **unique** global maximum Q_u: $F(Q_u) > F(Q)$ for all $Q \neq Q_u$.



The agent will choose only a global maximum as the quantity supplied to the market. However, it is possible that there are more than one global maximum. In that case, the supply quantity is not unique. Therefore, we prefer that the maximization problem has a unique global maximum.
A unique global maximum must be a global maximum and a global maximum must be a local maximum. $\Rightarrow$ to find a global maximum, we first find all the local maxima. One of them must be a global maximum, otherwise the problem does not have a solution (the maximum occurs at ∞.) We will find conditions (eg., increasing MC or decreasing MRS) so that there is only one local maximum which is also the unique global maximum.

7.3 FOC and SOC for a local maximum

At a local maximum Q_l, the slope of the graph of $F(Q)$ must be horizontal $F'(Q_l) = 0$. This is called the first order condition (FOC) for a local maximum.
A **critical** point Q_c: $F'(Q_c) = 0$.
A local maximum must be a critical point but a critical point does not have to be a local maximum.



A **degenerate** critical point: $F'(Q) = F''(Q) = 0$.
A non-degenerate critical point: $F'(Q) = 0$, $F''(Q) \neq 0$.
A non-degenerate critical point is a local maximum (minimum) if $F''(Q) < 0$ ($F''(Q) > 0$).

$$\text{FOC: } F'(Q_l) = 0 \qquad \text{SOC: } F''(Q_l) < 0$$

Example: $F(Q) = -15Q + 9Q^2 - Q^3$, $F'(Q) = -15 + 18Q - 3Q^2 = -3(Q-1)(Q-5)$. There are two critical points: $Q = 1, 5$. $F''(Q) = 18 - 6Q$, $F''(1) = 12 > 0$, and $F''(5) = -12 < 0$. Therefore, $Q = 5$ is a local maximum. It is a global maximum for $0 \leq Q < \infty$.

Remark 1 (Degeneracy): For a degenerate critical point, we have to check higher order derivatives. If the lowest order non-zero derivative is of odd order, then it is a reflect point; eg., $F(Q) = (Q-5)^3$, $F'(5) = F''(5) = 0$ and $F'''(5) = 6 \neq 0$ and $Q = 5$ is not a local maximum. If the lowest order non-zero derivative is of even order and negative (positive), then it is a local maximum (minimum); eg., $F(Q) = -(Q-5)^4$, $F'(5) = F''(5) = F'''(5) = 0$, $F^{(4)}(5) = -24 < 0$ and $Q = 5$ is a local maximum.
Remark 2 (Unboundedness): If $\lim_{Q\to\infty} F(Q) = \infty$, then a global maximum does not exist.
Remark 3 (Non-differentiability): If $F(Q)$ is not differentiable, then we have to use other methods to find a global maximum.
Remark 4 (Boundary or corner solution): When there is non-negative restriction $Q \geq 0$ (or an upper limit $Q \leq a$), it is possible that the solution occurs at $Q = 0$ (or at $Q = a$). To take care of such possibilities, FOC is modified to become $F'(Q) \leq 0$, $QF'(Q) = 0$ (or $F'(Q) \geq 0$, $(a-Q)F'(Q) = 0$).

7.4 Supply function of a competitive producer

Consider first the profit maximization problem of a competitive producer:

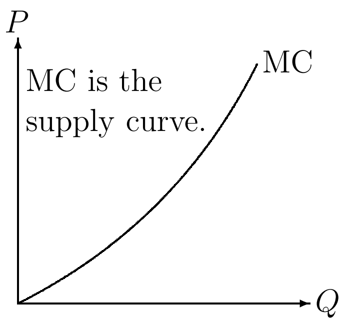
$$\max_Q \Pi = PQ - C(Q), \quad \text{FOC } \Rightarrow \frac{\partial \Pi}{\partial Q} = P - C'(Q) = 0.$$

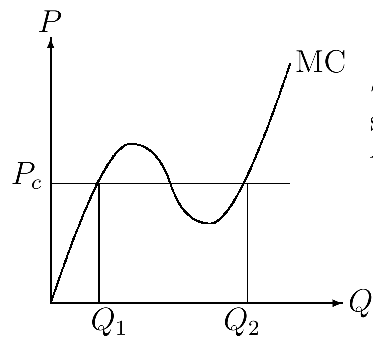
The FOC is the inverse supply function (a behavioral equation) of the producer: $P = C'(Q) = \text{MC}$. Remember that Q is endogenous and P is exogenous here. To find the comparative statics $\frac{dQ}{dP}$, we use the total differential method discussed in the last chapter:

$$dP = C''(Q)dQ, \;\; \Rightarrow \; \frac{dQ}{dP} = \frac{1}{C''(Q)}.$$

To determine the sign of $\frac{dQ}{dP}$, we need the SOC, which is $\frac{\partial^2 \Pi}{\partial Q^2} = -C''(Q) < 0$. Therefore, $\frac{dQ_s}{dP} > 0$.

Remark: The result is true no matter what the cost function $C(Q)$ is. MC $= C'(Q)$ can be non-monotonic, but the supply curve is only part of the increasing sections of the MC curve and can be discontinuous.





7.5 Maximization and comparative statics: general procedure

Maximization problem of an agent: $\max_X F(X;Y)$.
FOC: $F_X(X^*;Y) = 0, \Rightarrow X^* = X(Y) \cdots\cdots$ Behavioral Equation
Comparative statics: $F_{XX}dX + F_{XY}dY = 0 \Rightarrow \frac{dX}{dY} = -\frac{F_{XY}}{F_{XX}}$. SOC: $F_{XX} < 0$

Case 1: $F_{XY} > 0 \Rightarrow \frac{dX}{dY} = -\frac{F_{XY}}{F_{XX}} > 0$.
Case 2: $F_{XY} < 0 \Rightarrow \frac{dX}{dY} = -\frac{F_{XY}}{F_{XX}} < 0$.
Therefore, the sign of $\frac{dX}{dY}$ depends only on the sign of F_{XY}.

7.6 Utility Function

A consumer wants to maximize his/her utility function $U = u(Q) + M = u(Q) + (Y - PQ)$.
FOC: $\frac{\partial U}{\partial Q} = u'(Q) - P = 0$,
$\Rightarrow u'(Q_d) = P$ (inverse demand function)
$\Rightarrow Q_d = D(P)$ (demand function, a behavioral equation)

$\dfrac{\partial^2 U}{\partial Q \partial P} = U_{PQ} = -1 \Rightarrow \dfrac{dQ_d}{dP} = D'(P) < 0$, the demand function is a decreasing function of price.

7.7 Input Demand Function

The production function of a producer is given by $Q = f(x)$, where x is the quantity of an input employed. Its profit is $\Pi = pf(x) - wx$, where p (w) is the price of the output (input).
The FOC of profit maximization problem is $pf'(x) - w = 0$
$\Rightarrow f'(x) = w/p$ (inverse input demand function)
$\Rightarrow x = h(w/p)$ (input demand function, a behavioral equation)
$\dfrac{\partial^2 \Pi}{\partial x \partial (w/p)} = -1 \Rightarrow \dfrac{dx}{d(w/p)} = h'(w/p) < 0$, the input demand is a decreasing function of the real input price $\dfrac{w}{p}$.

7.8 Envelope theorem

Define the maximum function $M(Y) \equiv \max_X F(X,Y) = F(X(Y),Y)$ then the total derivative

$$\frac{dM}{dY} = M'(Y) = \left. \frac{\partial F(X,Y)}{\partial Y} \right|_{X=X(Y)}.$$

Proof: $M'(Y) = F_X \dfrac{dX}{dY} + F_Y$. At the maximization point $X = X(Y)$, FOC implies that the indirect effect of Y on M is zero.

In the consumer utility maximization problem, $V(P) \equiv U(D(P)) + Y - PD(P)$ is called the indirect utility function. The envelope theorem implies that $V'(P) = \left. \dfrac{dU}{dP} \right|_{Q_d = D(P)} \equiv -D(P)$, this is a simplified version of Roy's identity.

In the input demand function problem, $\pi(w,p) \equiv pf(h(w/p)) - wh(w/p)$ is the profit function. Let $p = 1$ and still write $\pi(w) \equiv f(h(w)) - wh(w)$. The envelope theorem implies that $\pi'(w) = \left. \dfrac{d\Pi}{dw} \right|_{x=h(w)} = -h(w)$, a simplified version of Hotelling's lemma.

Example: The relationships between LR and SR cost curves
STC$(Q;K) = C(Q,K)$, K: firm size. Each K corresponds to a STC.
LTC$(Q) = \min_K C(Q,K)$, $\Rightarrow K = K(Q)$ is the optimal firm size.
LTC is the envelope of STC's. Each STC tangents to the LTC (STC = LTC) at the quantity Q such that $K = K(Q)$. Notice that the endogenous variable is K and the exogenous is Q here.
By envelope theorem, LMC$(Q) = d$LTC$(Q)/dQ = \partial C(Q,K(Q))/\partial Q =SMC(Q;K(Q))$.
That is, when $K = K(Q)$ is optimal for producing Q, SMC=LMC.

Since LAC(Q) =LTC(Q)/Q and SAC(Q) =STC(Q)/Q, LAC is the envelope of SAC's and each SAC tangents to the LAC (SAC = LAC) at the quantity Q such that $K = K(Q)$.

7.9 Effect of a Unit Tax on Monopoly Output (Samuelson)

Assumptions: a monopoly firm, $q = D(P) \iff P = f(q)$, (inverse functions), $C = C(q)$, t = unit tax

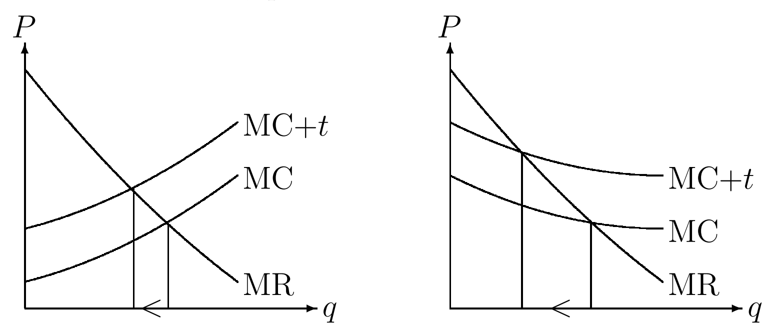
$$\max \pi(q) = Pq - C(q) - tq = qf(q) - C(q) - tq$$

q: endogenous; t: exogenous
FOC: $\partial\pi/\partial q = f(q) + qf'(q) - C'(q) - t = 0$.
The FOC defines a relationship between the monopoly output q^* and the tax rate t as $q^* = q(t)$ (a behavioral equation). The derivative dq^*/dt can be determined by the sign of the cross derivative:

$$\frac{\partial^2\pi}{\partial q \partial t} = -1 < 0$$

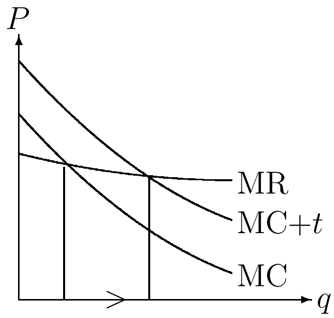
Therefore, we have $dq^*/dt < 0$.
The result can be obtained using the q–$p$diagram. FOC $\iff$ MR = MC + t. Therefore, on q–p space, an equilibrium is determined by the intersection point of MR and MC + t curves.

Case 1: MR is downward sloping and MC is upward sloping:
When t increases, q^* decreases as seen from the left diagram below.



Case 2: Both MR and MC are downward sloping and MR is steeper. MC decreasing; MR decreasing more $\Rightarrow t \uparrow q \downarrow$.

Case 3: Both MR and MC are downward sloping, but MC is steeper. The diagram shows that $dq^*/dt > 0$. It is opposite to our comparative statics result. Why? MR = MC + t violates SOC < 0, therefore, the intersection of MR and MC + t is not a profit maximizing point.



7.10 A Price taker vs a price setter

A producer employs an input X to produce an output Q. The production function is $Q = rX$. The inverse demand function for Q is $P = a - Q$. The inverse supply function of X is $W = b + X$, where W is the price of X. The producer is the only seller of Q and only buyer of X.

There are two markets, Q (output) and X (input). We want to find the 2-market equilibrium, i.e., the equilibrium values of W, X, Q, and P. It depends on the producer's power in each market.

Case 1: The firm is a monopolist in Q and a price taker in X. To the producer, P is endogenous and W is exogenous. Given W, its object is

$$\max_x \pi = PQ - WX = (a-Q)Q - WX = (a-rX)(rX) - WX, \quad \Rightarrow \text{FOC } ar - 2r^2X - W = 0.$$

The input demand function is $X = X(W) = \frac{ar-W}{2r^2}$.
Equating the demand and supply of X, the input market equilibrium is $X = \frac{ar-b}{2r^2+1}$ and $W = b + X = \frac{ar+2r^2b}{2r^2+1}$.
Substituting back into the production and output demand functions, the output market equilibrium is $Q = rX = far^2 - br2r^2 + 1$ and $P = a - Q = fa + ar^2 + b2r^2 + 1$.

Case 2: The firm is a price taker in Q and a monopsony in X. To the producer, P is exogenous and W is endogenous. Given P, its object is

$$\max_Q \pi = PQ - (b+X)X = PQ - (b + (Q/r))(Q/r), \quad \Rightarrow \text{FOC } P - \frac{b}{r} - \frac{2Q}{r^2} = 0.$$

The output supply function is $Q = Q(P) = \frac{r^2P-br}{2}$.
Equating the demand and supply of Q, the output market equilibrium is $Q = \frac{ar^2-br}{r^2-2}$ and $P = a - Q = \frac{2a+br}{2r^2+1}$.
Substituting back into the production and output demand functions, the output market equilibrium is $X = Q/r = far - br^2 - 2$ and $W = b + X = far + br^2 - 3br^2 - 2$.

Case 3: The firm is a monopolist in Q and a monopsony in X. To the producer, both P and W are endogenous, its object is

$$\max_x \pi = (a-Q)Q - (b+X)X = (a-(rX))(rX) - (b+X)X.$$

(We can also eliminate X instead of Q. The two procedures are the same.) (Show that π is strictly concave in X.) Find the profit maximizing X as a function of a and b, $X = X(a, b)$.
Determine the sign of the comparative statics $\dfrac{\partial X}{\partial a}$ and $\dfrac{\partial X}{\partial b}$ and explain your results in economics.
Derive the price and the wage rate set by the firm P and W and compare the results with that of cases 1 and 2.

7.11 Labor Supply Function

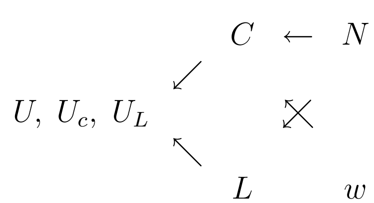
Consider a consumer/worker trying to maximize his utility function subject to the time constraint that he has only 168 hours a week to spend between work (N) and leisure (L), $N + L = 168$, and the budget constraint which equates his consumption (C) to his wage income (wN), $C = wN$, as follows:

$$\max U = U(C, L) = U(wN, 168 - N) \equiv f(N, w)$$

Here N is endogenous and w is exogenous. The FOC requires that the **total** derivative of U w.r.t. N be equal to 0.

$$\text{FOC: } \frac{dU}{dN} = f_N = U_C w + U_L(-1) = 0.$$

FOC defines a relationship between the labor supply of the consumer/worker, N, and the wage rate w, which is exactly the labor supply function of the individual $N^* = N(w)$. The slope the supply function $N'(w)$ is determined by the sign of the cross-derivative f_{Nw}



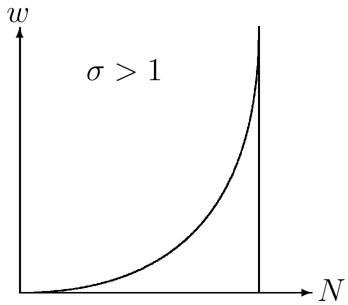
$$f_{Nw} = U_C + wNU_{CC} - NU_{LC}$$

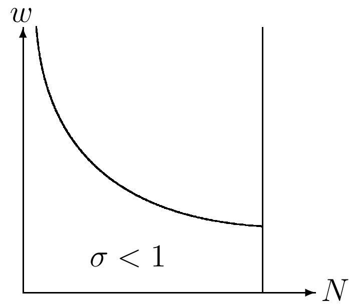
The sign of f_{Nw} is indeterminate, therefore, the slope of $N^* = N(w)$ is also indeterminate.
Numerical Examples:

Example 1: $U = 2\sqrt{C} + 2\sqrt{L}$ elasticity of substitution $\sigma > 1$

$$U = 2\sqrt{C}+2\sqrt{L} = 2\sqrt{wN}+2\sqrt{168-N}, \quad \frac{dU}{dN} = \frac{w}{\sqrt{C}} - \frac{1}{\sqrt{L}} = \frac{w}{\sqrt{wN}} - \frac{1}{\sqrt{168-N}} = 0.$$

Therefore, the inverse labor supply function is $w = N/(168 - N)$, which is positively sloped.





Example 2: $U = \frac{CL}{C+L}$ elasticity of substitution $\sigma < 1$

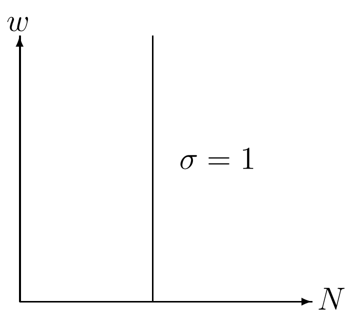
$$U = \frac{CL}{C+L} = \frac{wN(168-N)}{wN+168-N}, \quad \frac{dU}{dN} = \frac{wL^2 - C^2}{(C+L)^2} = 0.$$

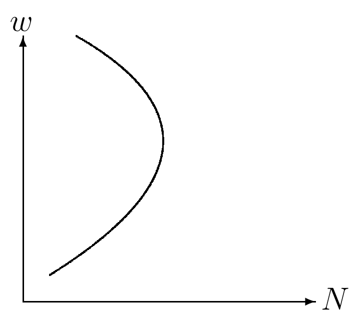
Therefore, the inverse labor supply function is $w = [(168 - N)/N]^2$, which is negatively sloped.

Example 3: $U = CL$ Cobb-Douglas($\sigma = 1$)

$$U = CL = wN(168 - N) \quad \frac{dU}{dN} = w(168 - 2N) = 0$$

The labor supply function is a constant $N = 84$ and the curve is vertical.





Backward-bending Labor Supply Curve: It is believed that the labor supply curve can be backward-bending.

7.12 Exponential and logarithmic functions and interest compounding

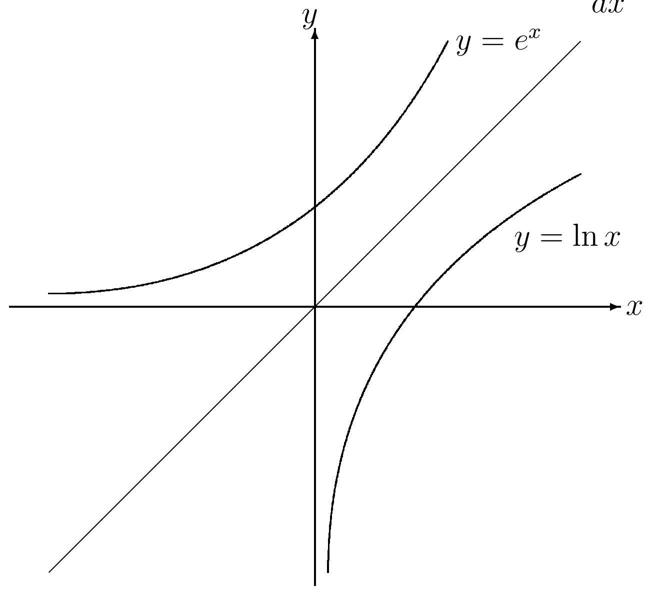
Exponential function $f(x) = e^x$ is characterized by $f(0) = 1$ and $f(x) = f'(x) = f''(x) = \ldots = f^{(n)}(x)$. The Taylor expansion of the exponential functionat $x = 0$ becomes

$$e^x = 1 + x + x^2/2! + \cdots + x^n/n! + \ldots$$

Some Rules: $\frac{d}{dx}(e^x) = e^x, \quad \frac{d}{dx}e^{ax} = ae^{ax}, \quad \frac{d}{dx}e^{f(x)} = f'(x)e^{f(x)}.$

The inverse of e^x is the logarithmic function: $\ln x$. If $x = e^y$, then we define $y \equiv \ln x$. More generally, if $x = a^y$, $a > 0$, then we define $y \equiv \log_a x = \dfrac{\ln x}{\ln a}$.

Using inverse function rule: $dx = e^y dy$, $\quad \dfrac{d}{dx}\ln x = \dfrac{1}{e^{\ln x}} = 1/x$.



$a^x = e^{\ln a^x} = e^{(\ln a)x} \Rightarrow \dfrac{d}{dx}a^x = (\ln a)e^{(\ln a)x} = (\ln a)a^x$.

$\dfrac{d}{dx}\ln_a x = \left(\dfrac{1}{\ln a}\right)\dfrac{1}{x}$

growth rate of a function of time $y = f(t)$:

$$\text{growth rate } \equiv \frac{1}{y}\frac{dy}{dt} = \frac{f'}{f} = \frac{d}{dt}[\ln f(t)].$$

Example: $f(t) = g(t)h(t)$. $\ln f(t) = \ln g(t) + \ln h(t)$, therefore, growth rate of $f(t)$ is equal to the sum of the growth rates of $g(t)$ and $h(t)$.

Interest compounding

A: principal (PV), V = Future Value, r = interest rate, n = number of periods

$$V = A(1+r)^n$$

If we compound interest m times per period, then the interest rate each time becomes r/m, the number of times of compounding becomes mn, and

$$V = A[(1 + r/m)^m]^n$$

$$\lim_{m\to\infty}(1 + r/m)^m = 1 + r + r^2/2! + \cdots + r^n/n! + \ldots = e^r$$

Therefore, $V \to Ae^{rn}$, this is the formula for instantaneous compounding.

7.13 Timing: (when to cut a tree)

t: number of years to wait, $A(t)$ present value after t years
$V(t) = Ke^{\sqrt{t}}$: the market value of the tree after t years
$A(t) = Ke^{\sqrt{t}}e^{-rt}$ is the present value
We want to find the optimal t such that the present value is maximized.

$$\max_t A(t) = Ke^{\sqrt{t}}e^{-rt}.$$

The FOC is

$$A' = \left(\frac{1}{2\sqrt{t}} - r\right) A(t) = 0$$

Because $A(t) \neq 0$, FOC implies: $\dfrac{1}{2\sqrt{t}} - r = 0$, $t^* = 1/(4r^2)$
For example, if $r = 10\%$, $t = 25$, then to wait 25 years before cutting the tree is the optimum.
Suppose that $A(t) = e^{f(t)}$
FOC becomes: $A'(t)/A(t) = f'(t) = r$, $\Rightarrow f'(t)$ is the instantaneous growth rate (or the marginal growth rate) at t, $\Rightarrow$ at $t = 25$, growth rate $= 10\% = r$, at $t = 26$, growth rate $< 10\%$. Therefore, it is better to cut and sell the tree at $t = 25$ and put the proceed in the bank than waiting longer.
SOC: It can be shown that $A'' < 0$.

7.14 Problems

1. Suppose the total cost function of a competitive firm is $C(Q) = e^{aQ+b}$. Derive the supply function.

2. The input demand function of a competitive producer, $X = X(W)$, can be derived by maximizing the profit function $\Pi = F(X) - WX$ with respect to X, where X is the quantity of input X and W is the price of X. Derive the comparative statics dX/dW and determine its sign.

3. The utility function of a consumer is given by $U = U(X) + M$, where X is the quantity of commodity X consumed and M is money. Suppose that the total income of the consumer is \$ 100 and that the price of X is P. Then the utility function become $U(X) + (100 - XP)$.

   (a) Find the first order condition of the utility maximization problem.

   (b) What is the behavior equation implied by the first order condition?

   (c) Derive dX/dP and determine its sign. What is the economic meaning of your result?

4. The consumption function of a consumer, $C = C(Y)$, can be derived by maximizing the utility function $U(C, Y) = u_1(C) + u_2(Y - C)$, where $u_1'(C) > 0$, $u_2'(Y - C) > 0$ and $u_1''(C) < 0$, $u_2''(Y - C) < 0$. Derive the comparative statics dC/dY and determine its sign.

5. Consider a duopoly market with two firms, A and B. The inverse demand function is $P = f(Q_A + Q_B)$, $f' < 0$, $f" < 0$. The cost function of firm A is $TC_A = C(Q_A)$, $C' > 0$, $C" > 0$. The profit of firm A is $\Pi_A = PQ_A - TC_A = Q_A f(Q_A + Q_B) - C(Q_A)$. For a given output of firm B, Q_B, there is a Q_A which maximizes firm A's profit. This relationship between Q_B and Q_A is called the reaction function of firm A, $Q_A = R_A(Q_B)$.

   (a) Find the slope of the reaction function $R'_A = \frac{dQ_A}{dQ_B}$.

   (b) When Q_B increases, will firm A's output Q_A increase or decrease?

6. The profit of a monopolistic firm is given by $\Pi = R(x) - C(x, b)$, where x is output, b is the price of oil, $R(x)$ is total revenue, and $C(x, b)$ is the total cost function. For any given oil price b, there is an optimal

   output which maximizes profit, that is, the optimal output is a function of oil price, $x = x(b)$. Assume that $C_{bx} = \partial^2 C/\partial b \partial x > 0$, that is, an increase in oil price will increase marginal cost. Will an increase in oil price increase output, that is, is $dx/db > 0$?

7. Consider a monopsony who uses a single input, labor (L), for the production of a commodity (Q), which he sells in a perfect competition market. His production function is $Q = F(L)$, ($f'(L) > 0$). The labor supply function is $L = L(w)$, or more convenient for this problem, $w = L^{-1}(L) = W(L)$. Given the commodity price p, there is an optimal labor input which maximizes the monopsonist's total profit $\Pi = pf(L) - W(L)L$. In this problem, you are asked to derive the relation between L and p.

   (a) State the FOC and the SOC of the profit maximization problem.

   (b) Derive the comparative statics dL/dp and determine its sign.

8. Suppose that a union has a fixed supply of labor (L) to sell, that unemployed workers are paid unemployment insurance at a rate of \$$u$ per worker, and that the union wishes to maximize the sum of the wage bill plus the unemployment compensation $S = wD(w) + u(L - D(w))$, where w is wage per

   worker, $D(w)$ is labor demand function, and $D'(w) < 0$. Show that if u increases, then the union should set a higher w. (Hint: w is endogenous and u is exogenous.)

9. The production function of a competitive firm is given by $Q = F(L, K)$. where L is variable input and K is fixed input. The short run profit function is given by $\Pi = pQ - wL - r\overline{K}$, where p is output price, w is wage rate, and r is the rental rate on capital. In the short run, given the quantity of fixed input $\overline{K}$,there is a L which maximizes Π. Hence, the short run demand for L can be regarded as a function of K. Assume that $F_{LK} > 0$.

   (a) State the FOC and SOC of the profit maximization problem.

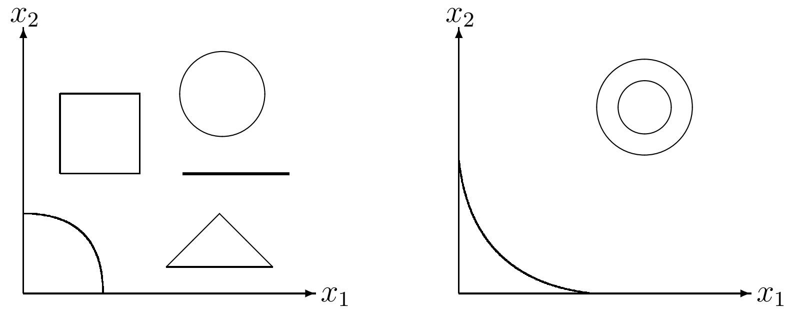
   (b) Derive the comparative statics dL/dK and determine its sign.

7.15 Concavity and Convexity

The derivation of a behavioral equation $X = X(Y)$ from the maximization problem $\max_x F(X;Y)$ is valid only if there exists a unique global maximum for every Y. If there are multiple global maximum, then $X = X(Y)$ has multiple values and the comparative static analysis is not valid. Here we are going to discuss a condition on $F(X;Y)$ so that a critical point is always a unique global maximum and the comparative static analysis is always valid.

Convex sets

A is a convex set if $\forall X^0, X^1 \in A$ and $0 \le \theta \le 1$, $X^\theta \equiv (1-\theta)X^0 + \theta X^1 \in A$. (If $X^0, X^1 \in A$ then the whole line connecting X^0 and X^1 is in A.)



some convex sets   some non-convex sets

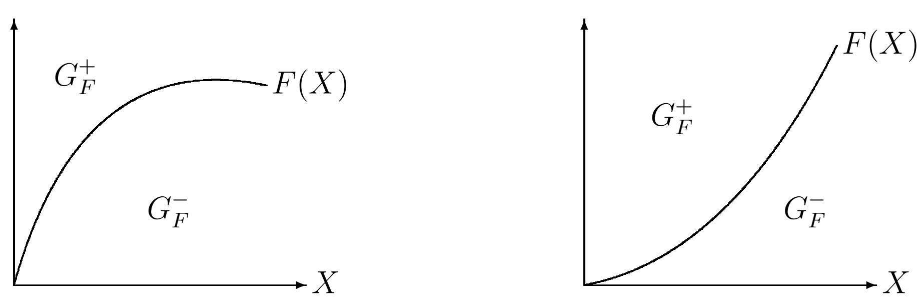
(1) If A_1 and A_2 are convex, then $A_1 \cap A_2$ is convex but $A_1 \cup A_2$ is not necessarily convex. Also, the empty set itself is a convex set.
(2) The convex hull of A is the smallest convex set that contains A. For example, the convex hull of $\{X^0\} \cup \{X^1\}$ is the straight line connecting X^0 and X^1.

Convex and concave functions

Given a function $F(X)$, we define the sets

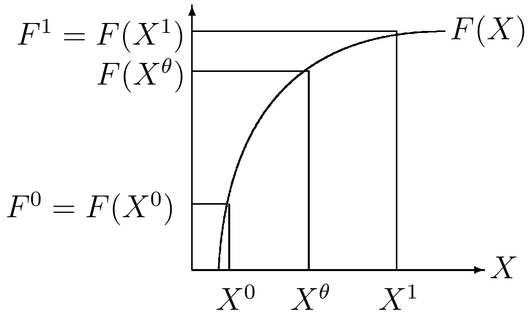
$$G_F^+ \equiv \{(x,y)|\, y \ge F(x), x \in R\}, \quad G_F^- \equiv \{(x,y)|\, y \le F(x), x \in R\}, \quad G_F^+,\ G_F^- \subset R^2.$$

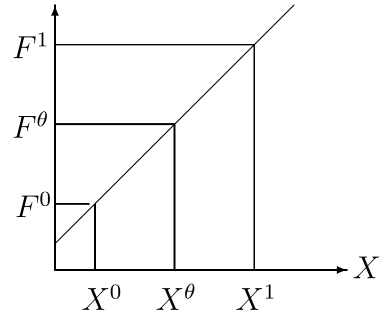
If G_F^+ (G_F^-) is a convex set, then we say $F(X)$ is a convex function (a concave function). If $F(X)$ is defined only for nonnegative values $X \ge 0$, the definition is similar.



G_F^- is a convex set $\Rightarrow F(X)$ is concave G_F^+ is a convex set $\Rightarrow F(X)$ is convex

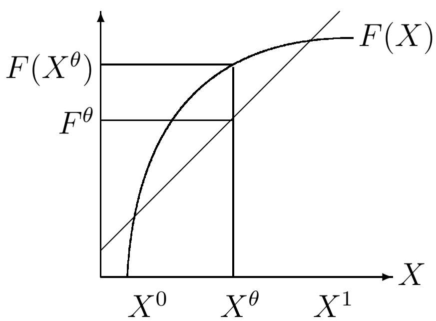
Equivalent Definition: Given $X^0 < X^1$, $0 \le \theta \le 1$, denote $F^0 = F(X^0)$, $F^1 = F(X^1)$. Define $X^\theta \equiv (1-\theta)X^0 + \theta X^1$, $F(X^\theta) = F((1-\theta)X^0 + \theta X^1)$. Also define $F^\theta \equiv (1-\theta)F(X^0) + \theta F(X^1) = (1-\theta)F^0 + \theta F^1$.



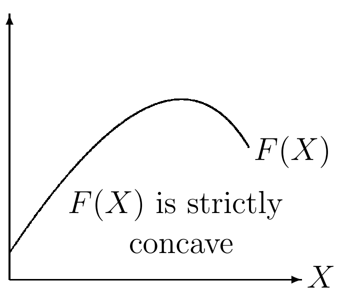


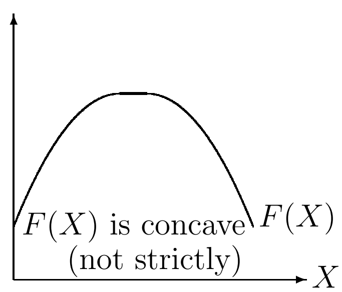
$$\frac{X^\theta - X^0}{X^1 - X^\theta} = \frac{\theta(X^1 - X^0)}{(1-\theta)(X^1 - X^0)} = \frac{\theta}{1-\theta}, \qquad \frac{F^\theta - F^0}{F^1 - F^\theta} = \frac{\theta(F^1 - F^0)}{(1-\theta)(F^1 - F^0)} = \frac{\theta}{1-\theta}.$$

Therefore, (X^θ, F^θ) is located on the straight line connecting (X^0, F^0) and (X^1, F^1) and when θ shifts from 0 to 1, (X^θ, F^θ) shifts from (X^0, F^0) to (X^1, F^1) (the right figure). On the other hand, $(X^\theta, F(X^\theta))$ shifts along the curve representing the graph of $F(X)$ (the left figure). Put the two figures together:



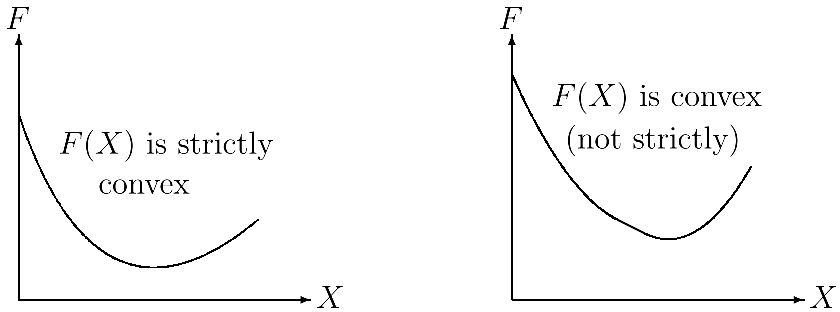
$F(X)$ is **strictly concave** $\Rightarrow$ if for all X^0, X^1 and $\theta \in (0,1)$, $F(X^\theta) > F^\theta$.
$F(X)$ is **concave** $\Rightarrow$ if for all X^0, X^1 and $\theta \in [0,1]$, $F(X^\theta) \geq F^\theta$.



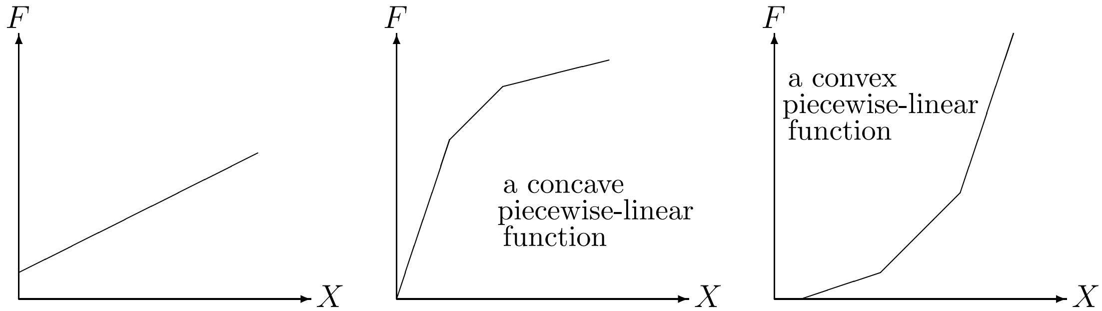


Notice that these concepts are global concepts. (They have something to do with the whole graph of F, not just the behavior of F nearby a point.) The graph of a concave function can have a flat part. For a strictly concave function, the graph should be curved everywhere except at kink points.

$F(X)$ is **strictly convex** $\Rightarrow$ if for all X^0, X^1 and $\theta \in (0,1)$, $F(X^\theta) < F^\theta$.
$F(X)$ is **convex** $\Rightarrow$ if for all X^0, X^1 and $\theta \in [0,1]$, $F(X^\theta) \leq F^\theta$.



Remark 1: A linear function is both concave and convex since $F^\theta \equiv F(X^\theta)$.



Remark 2: A piecewise-linear function consists of linear components; for example, the income tax schedule $T = f(Y)$ is a piecewise-linear function. Other examples are

$$\text{concave} \quad F(X) = \begin{cases} 2X & X \leq 1 \\ 1 + X & X > 1 \end{cases} \qquad \text{convex} \quad F(X) = \begin{cases} X & X \leq 1 \\ 2X - 1 & X > 1 \end{cases}$$

In the following theorems, we assume that $F''(X)$ exists for all X.

Theorem 1: $F(X)$ is concave, $\Leftrightarrow$ $F''(X) \leq 0$ for all X.
$F''(X) < 0$ for all X $\Rightarrow$ $F(X)$ is strictly concave.
Proof: By Taylor's theorem, there exist $\bar{X}^0 \in [X^0, X^\theta]$ and $\bar{X}^1 \in [X^\theta, X^1]$ such that

$$F(X^1) = F(X^\theta) + F'(X^\theta)(X^1 - X^\theta) + \frac{1}{2}F''(\bar{X}^1)(X^1 - X^\theta)^2$$

$$F(X^0) = F(X^\theta) + F'(X^\theta)(X^0 - X^\theta) + \frac{1}{2}F''(\bar{X}^0)(X^0 - X^\theta)^2$$

$$\Rightarrow \quad F^\theta = F(X^\theta) + \frac{1}{2}\theta(1-\theta)(X^1 - X^0)^2[F''(\bar{X}^0) + F''(\bar{X}^1)].$$

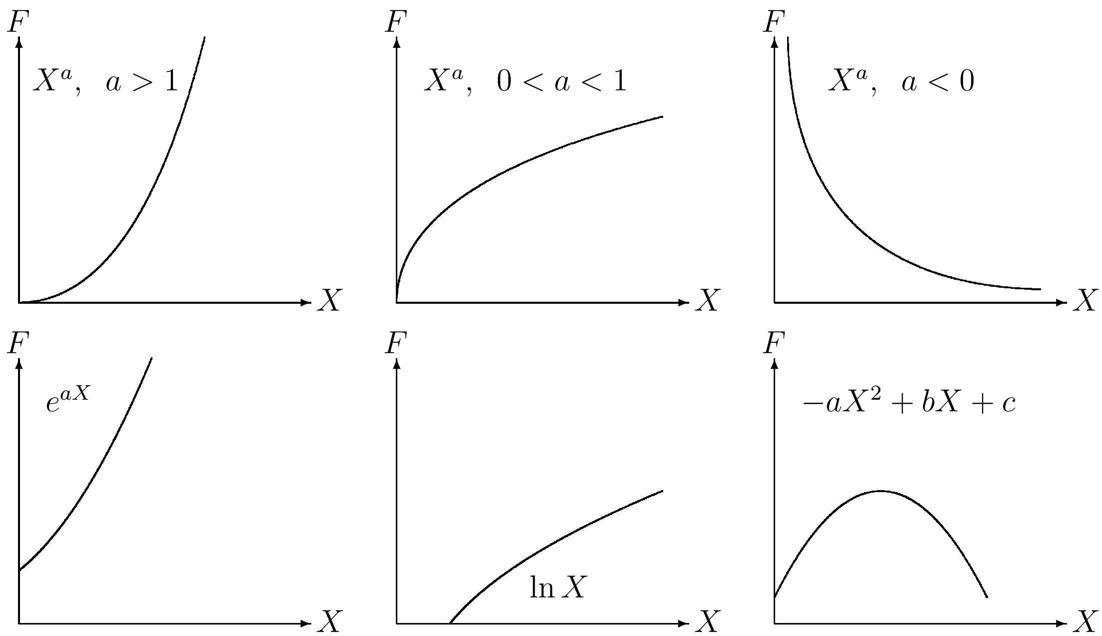
Theorem 2: If $F(X)$ is concave and $F'(X^0) = 0$, then X^0 is a global maximum.
If $F(X)$ is strictly concave and $F'(X^0) = 0$, then X^0 is a unique global maximum.
Proof: By theorem 1, X^0 must be a local maximum. If it is not a global maximum, then there exists X^1 such that $F(X^1) > F(X^0)$, which implies that $F(X^\theta) > F(X^0)$ for θ close to 0. Therefore, X^0 is not a local maximum, a contradiction.

Remark 1 (boundary/corner solution): The boundary or corner condition $F'(X) \leq 0$,

$XF'(X) = 0$ (or $F'(X) \geq 0$, $(X - a)F'(X) = 0$) becomes sufficient for global maximum.

Remark 2 (minimization problem): For the minimization problem, we replace concavity with convexity and $F''(X) < 0$ with $F''(X) > 0$. If $F(X)$ is convex and $F'(X^*) = 0$, then X^* is a global minimum.
If $F(X)$ is strictly convex and $F'(X^*) = 0$, then X^* is a unique global minimum.

Remark 3: The sum of two concave functions is concave. The product of two concave function is not necesarily concave. X^a is strictly concave if $a < 1$, strictly convex if $a > 1$. e^X is strictly convex and $\ln X$ is strctly concave with $X > 0$.



Remark 4: A concave function does not have to be differentiable, but it must be continuous on the interior points.

7.16 Indeterminate forms and L'Hôpital's rule

Let $f(x) = \dfrac{g(x)}{h(x)}$, $g(a) = h(a) = 0$ and $g(x)$ and $h(x)$ be continuous at $x = a$. $f(a)$ is not defined because it is $\dfrac{0}{0}$. However, $\lim_{x \to a} f(x)$ can be calculated.

$$\lim_{x \to a} f(x) = \lim_{x \to a} \frac{g(x)}{h(x)} = \lim_{x \to a} \frac{g(x) - g(a)}{h(x) - h(a)} = \lim_{\Delta x \to 0} \frac{g(a + \Delta x) - g(a)}{h(a + \Delta x) - h(a)} = \frac{g'(a)}{h'(a)}.$$

The same procedure also works for the case with $g(a) = h(a) = \infty$.

Example 1: $f(x) = \dfrac{g(x)}{h(x)} = \dfrac{\ln[(a^x + 1)/2]}{x}$,

$g(0) = h(0) = 0$, $h'(x) = 1$, $g'(x) = \dfrac{(\ln a)a^x}{a^x + 1}$.

$\Rightarrow\ h'(0) = 1$ and $g'(0) = \dfrac{\ln a}{2}$, $\Rightarrow\ \lim_{x\to 0} f(x) = \dfrac{\ln a}{2}$.

Example 2: $f(x) = \dfrac{g(x)}{h(x)} = \dfrac{x}{e^x}$, $h'(x) = e^x$, $g'(x) = 1$.

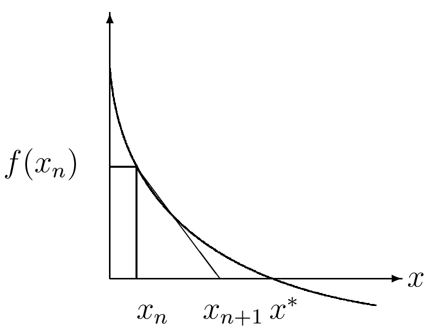
$\Rightarrow\ h'(\infty) = \infty$ and $g'(\infty) = 1$, $\Rightarrow\ \lim_{x\to\infty} f(x) = \dfrac{1}{\infty} = 0$.

Example 3: $f(x) = \dfrac{g(x)}{h(x)} = \dfrac{\ln x}{x}$, $h'(x) = 1$, $g'(x) = \dfrac{1}{x}$.

$\Rightarrow\ \lim_{x\to 0^+} h'(x) = 1$ and $\lim_{x\to 0^+} g'(x) = \infty$, $\Rightarrow\ \lim_{x\to 0^+} f(x) = \dfrac{1}{\infty} = 0$.

7.17 Newton's method

We can approximate a root x^* of a nonlinear equation $f(x) = 0$ using an algorithm called Newton's method.



$$f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}}$$

Recursive formula: $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$.

If $f(x)$ is not too strange, we will have $\lim_{n\to\infty} x_n = x^*$. Usually, two or three steps would be good enough.

Example: $f(x) = x^3 - 3$, $f'(x) = 3x^2$, $x_{n+1} = x_n - \dfrac{x^3 - 3}{3x^2} = x_n - \dfrac{x}{3} + \dfrac{1}{x^2}$.
Starting $x_0 = 1$, $x_1 = 1 - \dfrac{1}{3} + 1 = \dfrac{5}{3} \approx 1.666$. $x_2 = \dfrac{5}{3} - \dfrac{5}{9} + \dfrac{9}{25} = \dfrac{331}{225} \approx 1.47$. The true value is $x^* = \sqrt[3]{3} \approx 1.44225$.

8 Optimization–Multivariate Case

Suppose that the objective function has n variables: $\max_{x_1,\ldots,x_n} F(x_1,\ldots,x_n) = F(X)$.

A local maximum $X^* = (x_1^*,\ldots,x_n^*)$: $\exists \epsilon > 0$ such that $F(X^*) \geq F(X)$ for all X satisfying $x_i \in (x_i^* - \epsilon, x_i^* + \epsilon)\ \forall i$.
A critical point $X^c = (x_1^c,\ldots,x_n^c)$: $\dfrac{\partial F(X^c)}{\partial x_i} = 0\ \forall i$.
A global maximum X^*: $F(X^*) \geq F(X)\ \ \forall X$.
A unique global maximum X^*: $F(X^*) > F(X)\ \ \forall X \neq X^*$.
The procedure is the same as that of the single variable case. (1) Use FOC to find critical points; (2) use SOC to check whether a critical point is a local maximum, or show that $F(X)$ is concave so that a critical point must be a global maximum.

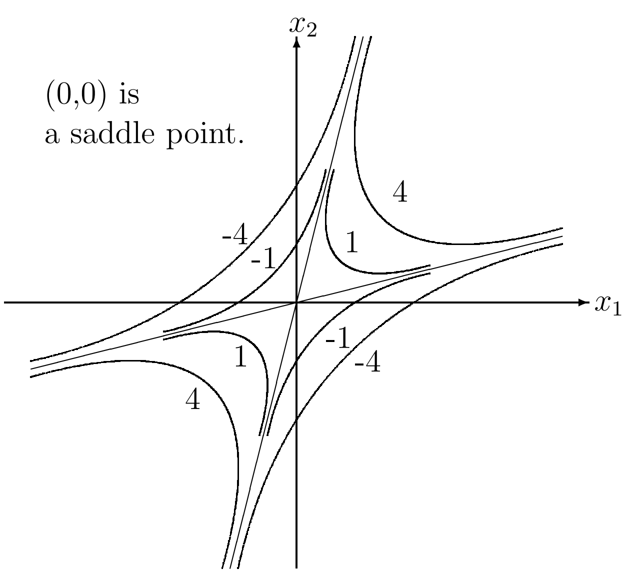
If we regard variables other than x_i as fixed, then it is a single variable maximization problem and, therefore, we have the necessary conditions

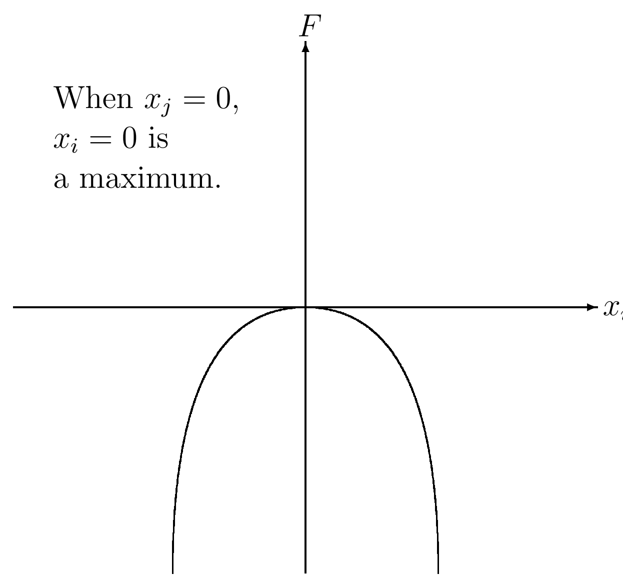
$$\text{FOC: } \frac{\partial F}{\partial x_i} = 0 \quad \text{and} \quad \text{SOC: } \frac{\partial^2 F}{\partial x_i^2} < 0, \quad i = 1,\ldots,n.$$

However, since there are $n \times n$ second order derivatives (the Hessian matrix $H(F)$) and the SOC above does not consider the cross-derivatives, we have a reason to suspect that the SOC is not sufficient. In the next section we will provide a counter-example and give a true SOC.

8.1 SOC

SOC of variable-wise maximization is wrong
Example: $\max_{x_1,x_2} F(x_1,x_2) = -x_1^2 + 4x_1x_2 - x_2^2$.
FOC: $F_1 = -2x_1 + 4x_2 = 0,\ \ F_2 = 4x_1 - 2x_2 = 0,\ \ \Rightarrow x_1 = x_2 = 0$.
SOC? $\quad F_{11} = F_{22} = -2 < 0$.





$F(0,0) = 0 < F(k,k) = 2k^2$ for all $k \neq 0$. $\Rightarrow (0,0)$ is not a local maximum!
The true SOC should take into consideration the possibility that when x_1 and x_2 increase simultaneously F may increase even if individual changes cannot increase F.

SOC of sequential maximization:
We can solve the maximization problem sequentially. That is, for a given x_2, we can find a x_1 such that $F(x_1, x_2)$ is maximized. The FOC and SOC are $F_1 = 0$ and $F_{11} < 0$. The solution depends on x_2, $x_1 = h(x_2)$, i.e., we regard x_1 as endogenous variable and x_2 as an exogenous variable in the first stage. Using implicit function rule, $dx_1/dx_2 = h'(x_2) = -F_{12}/F_{11}$. In the second stage we maximize $M(x_2) \equiv F(h(x_2), x_2)$. The FOC is $M'(x_2) = F_1 h' + F_2 = F_2 = 0$ (since $F_1 = 0$). The SOC is

$$M''(x_2) = F_1 h'' + F_{11}(h')^2 + 2F_{12}h' + F_{22} = (-F_{12}^2 + F_{11}F_{22})/F_{11} < 0.$$

Therefore, the true SOC is: $F_{11} < 0$ and $F_{11}F_{22} - F_{12}^2 > 0$.
For the n-variable case, the sequential argument is more complicated and we use Taylor's expansion.

SOC using Taylor's expansion:
To find the true SOC, we can also use Taylor's theorem to expand $F(X)$ around a critical point X^*:

$$F(X) = F(X^*) + F_1(X^*)(x_1 - x_1^*) + F_2(X^*)(x_2 - x_2^*)$$

$$+\frac{1}{2}(x_1 - x_1^*, x_2 - x_2^*)\begin{pmatrix} F_{11}(X^*) & F_{12}(X^*) \\ F_{21}(X^*) & F_{22}(X^*) \end{pmatrix}\begin{pmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{pmatrix} + \text{ higher order terms.}$$

Since X^* is a critical point, $F_1(X^*) = F_2(X^*) = 0$ and we have the approximation for X close to X^*:

$$F(X) - F(X^*) \approx \frac{1}{2}(x_1 - x_1^*, x_2 - x_2^*)\begin{pmatrix} F_{11}(X^*) & F_{12}(X^*) \\ F_{21}(X^*) & F_{22}(X^*) \end{pmatrix}\begin{pmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{pmatrix} \equiv \frac{1}{2}v'H^*v.$$

True SOC: $v'H^*v < 0$ for all $v \neq 0$.

In the next section, we will derive a systematic method to test the true SOC.

8.2 Quadratic forms and their signs

A quadratic form in n variables is a second degree homogenous function.

$$f(v_1, \ldots, v_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}v_i v_j = (v_1, \ldots, v_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \equiv v'Av.$$

Since $v_i v_j = v_j v_i$, we can assume that $a_{ij} = a_{ji}$ so that A is symmetrical.
Example 1: $(v_1 - v_2)^2 = v_1^2 - 2v_1v_2 + v_2^2 = (v_1, v_2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

Example 2: $2v_1v_2 = (v_1, v_2)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

Example 3: $x_1^2 + 2x_2^2 + 6x_1x_3 = (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

$v'Av$ is called **negative semidefinite** if $v'Av \le 0$ for all $v \in R^n$.
$v'Av$ is called **negative definite** if $v'Av < 0$ for all $v \neq 0$.
$v'Av$ is called **positive semidefinite** if $v'Av \ge 0$ for all $v \in R^n$.
$v'Av$ is called **positive definite** if $v'Av > 0$ for all $v \neq 0$.

(1) $-v_1^2 - 2v_2^2 < 0$, if $v_1 \neq 0$ or $v_2 \neq 0 \Rightarrow$ negative.
(2) $-(v_1 - v_2)^2 \le 0$, $= 0$ if $v_1 = v_2 \Rightarrow$ negative semidefinite.
(3) $(v_1 + v_2)^2 \ge 0$, $= 0$ if $v_1 = -v_2 \Rightarrow$ positive semidefinite.
(4) $v_1^2 + v_2^2 > 0$, if $v_1 \neq 0$ or $v_2 \neq 0 \Rightarrow$ positive.
(5) $v_1^2 - v_2^2 \overset{>}{\underset{<}{}} 0$ if $|v_1| \overset{>}{\underset{<}{}} |v_2| \Rightarrow$ neither positive nor negative definite.
(6) $v_1v_2 \overset{>}{\underset{<}{}} 0$ if $\mathrm{sign}(v_1) \overset{=}{\underset{\neq}{}} \mathrm{sign}(v_2) \Rightarrow$ neither positive nor negative definite.

Notice that if $v'Av$ is negative (positive) definite then it must be negative (positive) semidefinite.

Testing the sign of a quadratic form:

$n = 2$: $v'Av = (v_1, v_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = a_{11}v_1^2 + 2a_{12}v_1v_2 + a_{22}v_2^2 =$

$$a_{11}\left(v_1^2 + 2\frac{a_{12}}{a_{11}}v_1v_2 + \frac{a_{12}^2}{a_{11}^2}v_2^2\right) + \left(-\frac{a_{12}^2}{a_{11}} + a_{22}\right)v_2^2 = a_{11}\left(v_1 + \frac{a_{12}}{a_{11}}v_2\right)^2 + \frac{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{a_{11}}v_2^2.$$

Negative definite $\Leftrightarrow$ $a_{11} < 0$ and $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2 > 0$.

Positive definite $\Leftrightarrow$ $a_{11} > 0$ and $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2 > 0$.

For negative or positive semidefinite, replace strict inequalities with semi-inequalities. The proofs for them are more difficult and discussed in Lecture 13.

Example 1: $F(v_1, v_2) = -2v_1^2 + 8v_1v_2 - 2v_2^2 = (v_1, v_2)\begin{pmatrix} -2 & 4 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $a_{11} = -2 < 0$ but $\begin{vmatrix} -2 & 4 \\ 4 & -2 \end{vmatrix} = -12 < 0 \Rightarrow$ neither positive nor negative. The matrix is the Hessian of F of the counter-example in section 1. Therefore, the counter-example violates the SOC.

General n: $v'Av = (v_1, \ldots, v_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = a_{11}(v_1 + \cdots)^2$

$$+\frac{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{a_{11}}(v_2+\cdots)^2+\frac{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}(v_3+\cdots)^2+\cdots+\frac{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1(n-1)} \\ \vdots & \ddots & \vdots \\ a_{(n-1)1} & \cdots & a_{(n-1)(n-1)} \end{vmatrix}}v_n^2.$$

k-th order principle minor of A: $A^{(k)} \equiv \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, $k = 1, \ldots, n$.

Negative definite $\Leftrightarrow$ $A^{(1)} = a_{11} < 0$, $\dfrac{|A^{(2)}|}{|A^{(1)}|} < 0$, $\dfrac{|A^{(3)}|}{|A^{(2)}|} < 0$, ..., $\dfrac{|A^{(n)}|}{|A^{(n-1)}|} < 0$.

Positive definite $\Leftrightarrow$ $A^{(1)} = a_{11} > 0$, $\dfrac{|A^{(2)}|}{|A^{(1)}|} > 0$, $\dfrac{|A^{(3)}|}{|A^{(2)}|} > 0$, ..., $\dfrac{|A^{(n)}|}{|A^{(n-1)}|} > 0$.

Negative definite $\Leftrightarrow$ $|A^{(1)}| = a_{11} < 0$, $|A^{(2)}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$, $|A^{(3)}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0$, $\cdots$, $(-1)^n|A^{(n)}| = (-1)^n\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0$.

Positive definite $\Leftrightarrow$ $|A^{(1)}| = a_{11} > 0$, $|A^{(2)}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$, $|A^{(3)}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0$, $\cdots$, $|A^{(n)}| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0$.

The conditions for semidefinite are more complicated and will be discussed in Lecture 13 using the concept of eigenvalues of a square matrix.

8.3 SOC again

From the last two sections, the SOC for a local maximum can be summarized as:
SOC: $v'H_F(X^*)v$ is negative definite $\Rightarrow$ $F_{11}(X^*) < 0$, $\begin{vmatrix} F_{11}(X^*) & F_{12}(X^*) \\ F_{21}(X^*) & F_{22}(X^*) \end{vmatrix} = F_{11}F_{22} - F_{12}^2 > 0$, $\begin{vmatrix} F_{11}(X^*) & F_{12}(X^*) & F_{13}(X^*) \\ F_{21}(X^*) & F_{22}(X^*) & F_{23}(X^*) \\ F_{31}(X^*) & F_{32}(X^*) & F_{33}(X^*) \end{vmatrix} < 0$, ...

Example: $\max F(x_1, x_2) = 3x_1 + 3x_1x_2 - 3x_1^2 - x_2^3$. (A cubic function in 2 variables.)

FOC: $F_1 = 3 + 3x_2 - 6x_1 = 0$ and $F_2 = 3x_1 - 3x_2^2 = 0$.

Two critical points: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{2} \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Hessian matrix: $H(x_1, x_2) = \begin{pmatrix} F_{11}(X) & F_{12}(X) \\ F_{21}(X) & F_{22}(X) \end{pmatrix} = \begin{pmatrix} -6 & 3 \\ 3 & -6x_2 \end{pmatrix}$; $|H^1(X)| = F_{11}(X) = -6 < 0$, $|H^2(X)| = \begin{vmatrix} -6 & 3 \\ 3 & -6x_2 \end{vmatrix} = 36x_2 - 9$.

$|H^2(\frac{1}{4}, -\frac{1}{2})| = -27 < 0 \Rightarrow \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{2} \end{pmatrix}$ is not a local max.

$|H^2(1,1)| = 27 > 0 \Rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a local max. $F(1,1) = 2$. It is not a global maximum because $F \rightarrow \infty$ when $x_2 \rightarrow -\infty$.

Remark 1: $F_{11} < 0$ and $F_{11}F_{22} - F_{12}^2 > 0$ together implies $F_{22} < 0$.
Remark 2: If $|H^k(X^*)| = 0$ for some $1 \leq k \leq n$ then X^* is a degenerate critical point. Although we can still check whether $v'Hv$ is negative semidefinite, it is insufficient for a local maximum. We have to check the third or higher order derivatives to determine whether X^* is a local maximum. It is much more difficult than the single variable case with $F'' = 0$.

8.4 Joint products

A competitive producer produces two joint products. (eg., gasoline and its by products or cars and tructs, etc.)
Cost function: $C(q_1, q_2)$.
Profit function: $\Pi(q_1, q_2; p_1, p_2) = p_1 q_1 + p_2 q_2 - C(q_1, q_2)$.
FOC: $\Pi_1 = p_1 - C_1 = 0$, $\Pi_2 = p_2 - C_2 = 0$; or $p_i = \mathrm{MC}_i$.
SOC: $\Pi_{11} = -C_{11} < 0$, $\begin{vmatrix} -C_{11} & -C_{12} \\ -C_{21} & -C_{22} \end{vmatrix} \equiv \Delta > 0$.

Example: $C(q_1, q_2) = 2q_1^2 + 3q_2^2 - q_1 q_2$
FOC: $p_1 - 4q_1 + q_2 = 0$, $p_2 + q_1 - 6q_2 = 0 \Rightarrow$ supply functions $q_1 = (6p_1 + p_2)/23$ and $q_2 = (p_1 + 4p_2)/23$.
SOC: $-C_{11} = -4 < 0$, $\begin{vmatrix} -C_{11} & -C_{12} \\ -C_{21} & -C_{22} \end{vmatrix} = \begin{vmatrix} -4 & 1 \\ 1 & -6 \end{vmatrix} = 23 > 0$.

Comparative statics:
Total differential of FOC: $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} dq_1 \\ dq_2 \end{pmatrix} = \begin{pmatrix} dp_1 \\ dp_2 \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} \frac{\partial q_1}{\partial p_1} & \frac{\partial q_1}{\partial p_2} \\ \frac{\partial q_2}{\partial p_1} & \frac{\partial q_2}{\partial p_2} \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} C_{22} & -C_{12} \\ -C_{21} & C_{11} \end{pmatrix}.$$

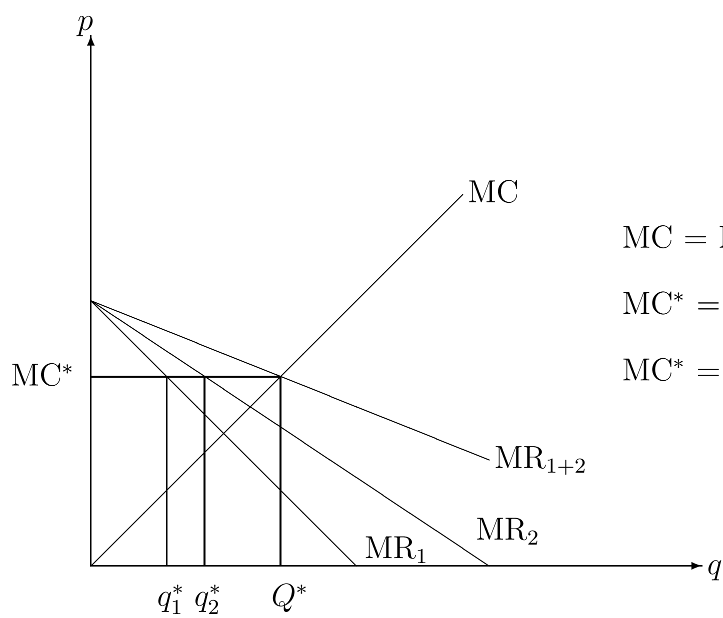
SOC $\Rightarrow C_{11} > 0,\ C_{22} > 0,\ \Delta > 0 \Rightarrow \dfrac{\partial q_1}{\partial p_1} > 0,\ \dfrac{\partial q_2}{\partial p_2} > 0.$

$\dfrac{\partial q_1}{\partial p_2}$ and $\dfrac{\partial q_2}{\partial p_1}$ are positive if the joint products are beneficially to each other in the production so that $C_{12} < 0$.

8.5 Monopoly price discrimination

A monopoly sells its product in two separable markets.
Cost function: $C(Q) = C(q_1 + q_2)$
Inverse market demands: $p_1 = f_1(q_1)$ and $p_2 = f_2(q_2)$
Profit function: $\Pi(q_1, q_2) = p_1q_1 + p_2q_2 - C(q_1 + q_2) = q_1f_1(q_1) + q_2f_2(q_2) - C(q_1 + q_2)$
FOC: $\Pi_1 = f_1(q_1) + q_1f_1'(q_1) - C'(q_1 + q_2) = 0$, $\Pi_2 = f_2(q_2) + q_2f_2'(q_2) - C'(q_1 + q_2) = 0$; or $\mathrm{MR}_1 = \mathrm{MR}_2 = \mathrm{MC}$.
SOC: $\Pi_{11} = 2f_1' + q_1f_1'' - C'' < 0$, $\begin{vmatrix} 2f_1' + q_1f_1'' - C'' & -C'' \\ -C'' & 2f_2' + q_2f_2'' - C'' \end{vmatrix} \equiv \Delta > 0.$

Example: $f_1 = a - bq_1,\ f_2 = \alpha - \beta q_2$, and $C(Q) = 0.5Q^2 = 0.5(q_1 + q_2)^2$.
$f_1' = -b,\ f_2' = -\beta,\ f_1'' = f_2'' = 0,\ C' = Q = q_1 + q_2$, and $C'' = 1$.
FOC: $a - 2bq_1 = q_1 + q_2 = \alpha - 2\beta q_2 \Rightarrow \begin{pmatrix} 1+2b & 1 \\ 1 & 1+2\beta \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} a \\ \alpha \end{pmatrix}$
$\Rightarrow \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \dfrac{1}{(1+2b)(1+2\beta) - 1}\begin{pmatrix} a(1+2\beta) - \alpha \\ \alpha(1+2b) - a \end{pmatrix}.$
SOC: $-2b - 1 < 0$ and $\Delta = (1+2b)(1+2\beta) - 1 > 0$.



$\mathrm{MC} = \mathrm{MR}_{1+2} \Rightarrow Q^*, \mathrm{MC}^*$

$\mathrm{MC}^* = \mathrm{MR}_1 \Rightarrow q_1^*$

$\mathrm{MC}^* = \mathrm{MR}_2 \Rightarrow q_2^*$

8.6 SR supply vs LR supply - Le Châtelier principle

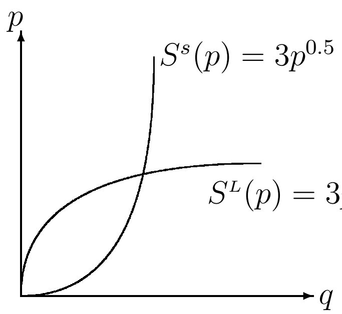
A competitive producer employs a variable input x_1 and a fixed input x_2. Assume that the input prices are both equal to 1, $w_1 = w_2 = 1$.
Production function: $q = f(x_1, x_2)$, assume $\mathrm{MP}_i = f_i > 0,\ f_{ii} < 0,\ f_{ij} > 0,$

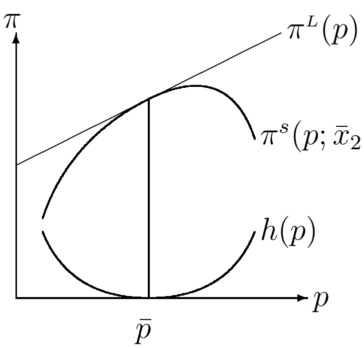
$f_{11}f_{22} > f_{12}^2$.
Profit function: $\Pi(x_1, x_2; p) = pf(x_1, x_2) - x_1 - x_2$.

Short-run problem (only x_1 can be adjusted, x_2 is fixed):
SR FOC: $\Pi_1 = pf_1 - 1 = 0$, or $w_1 = \text{VMP}_1$. SOC: $\Pi_{11} = pf_{11} < 0$,
Comparative statics: $f_1 dp + pf_{11} dx_1 = 0 \Rightarrow \dfrac{dx_1}{dp} = \dfrac{-f_1}{pf_{11}} > 0$, $\dfrac{dq^s}{dp} = \dfrac{-f_1^2}{pf_{11}} = \dfrac{-1}{p^3 f_{11}} > 0$.

Long-run problem (both x_1 and x_2 can be adjusted):
LR FOC: $\Pi_1 = pf_1 - 1 = 0$, $\Pi_2 = pf_2 - 1 = 0$; or $w_i = \text{VMP}_i$.
SOC: $\Pi_{11} = pf_{11} < 0$, $\begin{vmatrix} pf_{11} & pf_{12} \\ pf_{21} & pf_{22} \end{vmatrix} \equiv \Delta > 0$.
Comparative statics: $\begin{pmatrix} f_1 dp \\ f_2 dp \end{pmatrix} + \begin{pmatrix} pf_{11} & pf_{12} \\ pf_{21} & pf_{22} \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$\Rightarrow \begin{pmatrix} dx_1/dp \\ dx_2/dp \end{pmatrix} = \dfrac{-1}{p(f_{11}f_{22} - f_{12}^2)} \begin{pmatrix} f_1 f_{22} - f_2 f_{12} \\ f_2 f_{11} - f_1 f_{21} \end{pmatrix}$, $\dfrac{dq^L}{dp} = \dfrac{-(f_{11} + f_{22} - 2f_{12})}{p^3(f_{11}f_{22} - f_{12}^2)}$
Le Châtelier principle: $\dfrac{dq^L}{dp} > \dfrac{dq^s}{dp}$.

Example: $f(x_1, x_2) = 3x_1^{1/3}x_2^{1/3}$ (homogenous of degree 2/3)
LR FOC: $px_1^{-2/3}x_2^{1/3} = 1 = px_1^{1/3}x_2^{-2/3} \Rightarrow x_1 = x_2 = p^3$ $q^L = 3p^2$.
SR FOC (assuming $\bar{x}_2 = 1$): $px_1^{-2/3} = 1 \Rightarrow x_1 = p^{3/2}$ $q^s = 3p^{1/2}$.
η^L (LR supply elasticity) $= 2 > \eta^s$ (SR supply elasticity) $= 1/2$.





Envelop theorem, Hotelling's lemma, and Le Châtelier principle
From the SR problem we first derive the SR variable input demand function $x_1 = x_1(p, \bar{x}_2)$.
Then the SR supply function is obtained by substituting into the production function $q^s = f(x_1(p, \bar{x}_2), \bar{x}_2) \equiv S^s(p; \bar{x}_2)$.
The SR profit function is $\pi^s(p, \bar{x}_2) = pq^s - x_1(p, \bar{x}_2) - \bar{x}_2$.
Hotelling's lemma: By envelop theorem, $\dfrac{\partial \pi^s}{\partial p} = S^s(p, \bar{x}_2)$.

From the LR problem we first derive the input demand functions $x_1 = x_1(p)$ and $x_2 = x_2(p)$.
Then the LR supply function is obtained by substituting into the production function

$q^L = f(x_1(p), x_2(p)) \equiv S^L(p)$.
The LR profit function is $\pi^L(p) = pq^L - x_1(p) - x_2(p)$.
(Also Hotelling's lemma) By envelop theorem, $\dfrac{\partial \pi^L}{\partial p} = S^L(p)$.
Notice that $\pi^L(p) = \pi^s(p; x_2(p))$.

Let $\bar{x}_2 = x_2(\bar{p})$, define $h(p) \equiv \pi^L(p) - \pi^s(p, \bar{x}_2)$. $h(p) \geq 0$ because in the LR, the producer can adjust x_2 to achieve a higher profit level.
Also, $h(\bar{p}) = 0$ because $\pi^L(\bar{p}) = \pi^s(\bar{p}; x_2(\bar{p})) = \pi^s(\bar{p}; \bar{x}_2)$.
Therefore, $h(p)$ has a minimum at $p = \bar{p}$ and the SOC is $h''(\bar{p}) > 0 = \dfrac{\partial^2 \pi^L}{\partial p^2} - \dfrac{\partial^2 \pi^s}{\partial p^2} > 0$, which implies
Le Châtelier principle: $\dfrac{dq^L}{dp} - \dfrac{dq^s}{dp} > 0$.

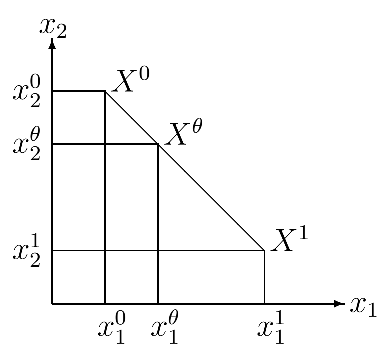
8.7 Concavity and Convexity

Similar to the single variable case, we define the concepts of concavity and convexity for 2-variable functions $F(X) = F(x_1, x_2)$ by defining $G_F^+,\ G_F^- \subset R^2$ s follows.

$$G_F^+ \equiv \{(x_1, x_2, y) |\ y \geq F(x_1, x_2), (x_1, x_2) \in R^2\}, \quad G_F^- \equiv \{(x_1, x_2, y) |\ y \leq F(x_1, x_2), (x_1, x_2) \in R^2\}.$$

If G_F^+ (G_F^-) is a convex set, then we say $F(X)$ is a convex function (a concave function). If $F(X)$ is defined only for nonnegative values $x_1, x_2 \geq 0$, the definition is similar. (The extension to n-variable functions is similar.)

Equivalently, given $X^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix}$, $X^1 = \begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix}$, $0 \leq \theta \leq 1$, $F^0 \equiv F(X^0)$, $F^1 \equiv F(X^1)$, we define
$X^\theta \equiv (1-\theta)X^0 + \theta X^1 = \begin{pmatrix} (1-\theta)x_1^0 + \theta x_1^1 \\ (1-\theta)x_2^0 + \theta x_2^1 \end{pmatrix} \equiv \begin{pmatrix} x_1^\theta \\ x_2^\theta \end{pmatrix}$, $F(X^\theta) = F((1-\theta)X^0 + \theta X^1)$
$F^\theta \equiv (1-\theta)F(X^0) + \theta F(X^1) = (1-\theta)F^0 + \theta F^1$.



X^θ is located on the straight line connecting X^0 and X^1, when θ shifts from 0 to 1, X^θ shifts from X^0 to X^1.

On 3-dimensional (x_1–x_2–F) space, (X^θ, F^θ) is located on the straight line connecting (X^0, F^0) and (X^1, F^1), when θ shifts from 0 to 1, (X^θ, F^θ) shifts from (X^0, F^0) to (X^1, F^1). On the other hand, $(X^\theta, F(X^\theta))$ shifts along the surface representing the graph of $F(X)$.

$F(X)$ is **strictly concave** $\Rightarrow$ if for all X^0, X^1 and $\theta \in (0,1)$, $F(X^\theta) > F^\theta$.
$F(X)$ is **concave** $\Rightarrow$ if for all X^0, X^1 and $\theta \in [0,1]$, $F(X^\theta) \geq F^\theta$.
$F(X)$ is **strictly convex** $\Rightarrow$ if for all X^0, X^1 and $\theta \in (0,1)$, $F(X^\theta) < F^\theta$.
$F(X)$ is **convex** $\Rightarrow$ if for all X^0, X^1 and $\theta \in [0,1]$, $F(X^\theta) \leq F^\theta$.

Example: $9x_1^{1/3}x_2^{1/3}$.

Assume that F is twice differentiable.
Theorem 1: $F(X)$ is concave, $\Leftrightarrow$ $v'H_Fv$ is negative semidefinite for all X.
$v'H_Fv$ is negative definite for all X $\Rightarrow$ $F(X)$ is strictly concave.
Proof: By Taylor's theorem, there exist $\bar{X}^0 \in [X^0, X^\theta]$ and $\bar{X}^1 \in [X^\theta, X^1]$ such that

$$F(X^1) = F(X^\theta) + \nabla F(X^\theta)(X^1 - X^\theta) + \frac{1}{2}(X^1 - X^\theta)'H_F(\bar{X}^1)(X^1 - X^\theta)$$

$$F(X^0) = F(X^\theta) + \nabla F(X^\theta)(X^0 - X^\theta) + \frac{1}{2}(X^0 - X^\theta)'H_F(\bar{X}^0)(X^0 - X^\theta)$$

$$\Rightarrow \quad F^\theta = F(X^\theta) + \frac{\theta(1-\theta)}{2}(X^1 - X^0)'[H_F(\bar{X}^0) + H_F(\bar{X}^1)](X^1 - X^0).$$

Theorem 2: If $F(X)$ is concave and $\nabla F(X^0) = 0$, then X^0 is a global maximum.
If $F(X)$ is strictly concave and $\nabla F(X^0) = 0$, then X^0 is a unique global maximum.
Proof: By theorem 1, X^0 must be a local maximum. If it is not a global maximum, then there exists X^1 such that $F(X^1) > F(X^0)$, which implies that $F(X^\theta) > F(X^0)$ for θ close to 0. Therefore, X^0 is not a local maximum, a contradiction.

Remark 1 (boundary or corner solution): The boundary or corner condition $F_i(X) \leq 0$, $x_iF_i(X) = 0$ (or $F_i(X) \geq 0$, $(x_i - a_i)F_i(X) = 0$) becomes sufficient for global maximum.
Remark 2 (minimization problem): For the minimization problem, we replace concavity with convexity and negative definite with positive definite. If $F(X)$ is convex and $\nabla F(X^*) = 0$, then X^* is a global minimum.
If $F(X)$ is strictly convex and $\nabla F(X^*) = 0$, then X^* is a unique global minimum.

8.8 Learning and utility maximization

Consumer B's utility function is

$$U = u(x,k) + m - h(k), \quad x, m, k \geq 0,$$

where x is the quantity of commodity X consumed, k is B's knowledge regarding the consumption of X, m is money, $u(x,k)$ is the utility obtained, $\frac{\partial^2 u}{\partial x^2} < 0$, $\frac{\partial^2 u}{\partial x \partial k} > 0$ (marginal utility of X increases with k), and $h(k)$ is the disutility of acquiring/maintaining k, $h' > 0$, $h'' > 0$. Assume that B has 100 dollar to spend and the price of X is $P_x = 1$ so that $m = 100 - x$ and $U = u(x,k) + 100 - x - h(k)$. Assume also that k is **fixed in the short run**. The short run utility maximization problem is

$$\max_x u(x,k) + 100 - x - h(k), \quad \Rightarrow \text{FOC: } \frac{\partial u}{\partial x} - 1 = 0, \quad \text{SOC: } \frac{\partial^2 u}{\partial x^2} < 0.$$

The short run comparative static $\frac{dx}{dk}$ is derived from FOC as

$$\frac{dx}{dk} = -\frac{\partial^2 u/\partial x^2}{\partial^2 u/\partial x \partial k} > 0,$$

that is, consumer B will consume more of X if B's knowledge of X increases.

In the long run consumer B will change k to attain higher utility level. The long run utility maximization problem is

$$\max_{x,k} u(x,k) + 100 - x - h(k), \quad \Rightarrow \text{FOC: } \frac{\partial u}{\partial x} - 1 = 0, \quad \frac{\partial u}{\partial k} - h'(k) = 0.$$

The SOC is satisfied if we assume that $u(x,k)$ is concave and $h(k)$ is convex ($h''(k) > 0$), because $F(x,k) \equiv u(x,k) + 100 - x - h(k)$, $x, k \geq 0$ is strictly concave then.

Consider now the specific case when

$$u(x,k) = 3x^{2/3}k^{1/3} \quad \text{and} \quad h(k) = 0.5k^2.$$

1. Calculate consumer B's short run consumption of X, $x_s = x(k)$. (In this part, you may ignore the nonnegative constraint $m = 100 - x \geq 0$ and the possibility of a corner solution.)
   $x = 1/(8k)$.
2. Calculate the long run consumption of X, x_L.
   $x = 32$
3. Derive the short run value function $V(k) \equiv u(x(k),k) + 100 - x(k) - h(k)$.
4. Solve the maximization problem $\max_k V(k)$.
   $k = 4$.
5. Explain how the SOC is satisfied and why the solution is the unique global maximum.

(demand functions) Consider now the general case when $P_x = p$ so that

$$U = 3x^{2/3}k^{1/3} + 100 - px - 0.5k^2.$$

1. Calculate consumer B's short run demand function of X, $x_s = x(p;k)$. (Warning: There is a nonnegative constraint $m = 100 - px \geq 0$ and you have to consider both interior and corner cases.)
   $x_s(p) = 8kp^{-3}$ $(100/p)$ if $p \leq \sqrt{2k}/5$ $(p < \sqrt{2k}/5)$.
2. Calculate the long run demand function of X, $x_L(p)$. and the optimal level of K, $k_L(p)$. (Both interior and corner cases should be considered too.)
   $x_L(p) = 32p^{-5}$ $(100/p)$ if $p > (8/25)^{1/4}$ $(p < (8/25)^{1/4})$, $k_L(p) = [x_L(p)]^{2/5}$.

8.9 Homogeneous and homothetic functions

A homogeneous function of degree k is $f(x_1, \ldots, x_n)$ such that

$$f(hx_1, \ldots, hx_n) = h^k f(x_1, \ldots, x_n) \ \ \forall h > 0.$$

If f is homogeneous of degree k_1 and g homogeneous of degree k_2, then fg is homogeneous of degree $k_1 + k_2$, $\frac{f}{g}$ is homogeneous of degree $k_1 - k_2$, and f^m is homogeneous of degree mk_1.
If $Q = F(x_1, x_2)$ is homogeneous of degree 1, then $Q = x_1 F(1, \frac{x_2}{x_1}) \equiv x_1 f\left(\frac{x_2}{x_1}\right)$. If $m = H(x_1, x_2)$ is homogeneous of degree 0, then $m = H(1, \frac{x_2}{x_1}) \equiv h\left(\frac{x_2}{x_1}\right)$.
Euler theorem: If $f(x_1, \ldots, x_n)$ is homogeneous of degree k, then

$$x_1 f_1 + \ldots + x_n f_n = kf$$

If f is homogeneous of degree k, then f_i is homogenous of degree $k - 1$.

Examples: 1. Cobb-Douglas function $Q = Ax_1^{\alpha} x_2^{\beta}$ is homogeneous of degree $\alpha + \beta$.
2. CES function $Q = \{ax_1^{\rho} + bx_2^{\rho}\}^{\frac{k}{\rho}}$ is homogenous of degree k.
3. A quadratic form $x'Ax$ is homogeneous of degree 2.
4. Leontief function $Q = \min\{ax_1, bx_2\}$ is homogeneous of degree 1.

Homothetic functions: If f is homogeneous and $g = H(f)$, H is a monotonic increasing function, then g is called a homothetic function.

Example: $Q = \alpha \ln x_1 + \beta \ln x_2 = \ln\left(x_1^{\alpha} x_2^{\beta}\right)$ is homothetic.

The MRS of a homothetic function depends only on the ratio $\frac{x_2}{x_1}$.

8.10 Problems

1. Use Newton's method to find the root of the nonlinear equation $X^3 + 2X + 2 = 0$ accurate to 2 digits.

2. Find (a) $\lim_{X \to 0} \frac{1 - 2^{-X}}{X}$, (b) $\lim_{X \to 0^+} \frac{1 - e^{-aX}}{X}$, (c) $\lim_{X \to 0} \frac{e^{2X} - e^X}{X}$.

3. Given the total cost function $C(Q) = e^{aQ+b}$, use L'hôpital's rule to find the AVC at $Q = 0^+$.

4. Let $z = x_1x_2 + x_1^2 + 3x_2^2 + x_2x_3 + x_3^2$.
   (a) Use matrix multiplication to represent z.
   (b) Determine whether z is positive definite or negative definite.
   (c) Find the extreme value of z. Check whether it is a maximum or a minimum.

5. The cost function of a competitive producer is $C(Q;K) = \dfrac{Q^3}{3K} + K$, where K is, say, the plant size (a fixed factor in the short run).

   (a) At which output level, the SAC curve attains a minimum?

   (b) Suppose that the equilibrium price is $p = 100$. The profit is $\Pi = 100Q - \dfrac{Q^3}{3K} - K$. For a given K, find the supply quantity $Q(K)$ such that SR profit is maximized.

   (c) Calculate the SR maximizing profit $\pi(K)$.

   (d) Find the LR optimal $K = K^*$ to maximize $\pi(K)$.

   (e) Calculate the LR supply quantity $Q^* = Q(K^*)$.

   (f) Now solve the 2-variable maximization problem

   $$\max_{Q,K\geq 0} \ \Pi(Q,K) = pQ - C(Q) = 100Q - \frac{Q^3}{3K} - K.$$

   and show that $\Pi(Q,K)$ is concave so that the solution is the unique global maximum.

6. A competitive firm produces two joint products. The total cost function is $C(q_1,q_2) = 2q_1^2 + 3q_2^2 - 4q_1q_2$.
   (a) Use the first order conditions for profit maximization to derive the supply functions.
   (b) Check that the second order conditions are satisfied.

7. Check whether the function $f(x,y) = e^{x+y}$ is concave, convex, or neither.

8. (a) Check whether $f(x,y) = 2\ln x + 3\ln y - x - 2y$ is concave, convex, or neither. (Assume that $x > 0$ and $y > 0$.)
   (b) Find the critical point of f.
   (c) Is the critical point a local maximum, a global maximum, or neither?

9. Suppose that a monopoly can produce any level of output at a constant marginal cost of \$$c$ per unit. Assume that the monopoly sells its goods in two different markets which are separated by some distance. The demand curve in the first market is given by $Q_1 = \exp[-aP_1]$ and the curve in the second market is given by $Q_2 = \exp[-bP_2]$. If the monopolist wants to maximize its total profits, what level of output should be produced in each market and what price will prevail in each market? Check that your answer is the unique global maximum. (Hints: 1. $P_1 = -(1/a)\ln Q_1$ and $P_2 = -(1/b)\ln Q_2$. 2. $\Pi = P_1Q_1 + P_2Q_2 - (Q_1+Q_2)c = -(1/a)Q_1\ln Q_1 - (1/b)Q_2\ln Q_2 - (Q_1+Q_2)c$ is strictly concave.)

10. The production function of a competitive firm is given by $q = f(x_1,x_2) = 3x_1^{\frac{1}{3}}x_2^{\frac{1}{3}}$, where x_1 is a variable input and x_2 is a fixed input. Assume that the prices of the output and the fixed input are $p = w_2 = 1$. In the short run, the amount of the fixed input is given by $x_2 = \bar{x}_2$. The profit function of the competitive firm is given by $\pi = f(x_1,\bar{x}_2) - w_1x_1 - \bar{x}_2$.

(a) State the FOC for the SR profit maximization and calculate the SR input demand function $x_1^S = x_1^S(w_1)$ and the SR input demand elasticity $\frac{w_1}{x_1^S}\frac{\partial x_1^S}{\partial w_1}$.

(b) Now consider the LR situation when x_2 can be adjusted. State the FOC for LR profit maximization and calculate the LR input demand funtion $x_1^L = x_1^L(w_1)$ and the LR input demand elasticity $\frac{w_1}{x_1^L}\frac{\partial x_1^L}{\partial w_1}$.

(c) Verify the Le Châtelier principle: $\left|\frac{w_1}{x_1^L}\frac{\partial x_1^L}{\partial w_1}\right| > \left|\frac{w_1}{x_1^S}\frac{\partial x_1^S}{\partial w_1}\right|$.

(d) Show that the LR profit is a strictly concave function of (x_1, x_2) for $x_1, x_2 > 0$ and therefore the solution must be the unique global maximum.

11. Let $U(x, y) = x^a y + xy^2$, $x, y, a > 0$.

(a) For what value(s) of a $U(x, y)$ is homogeneous?.

(b) For what value(s) of a $U(x, y)$ is homothetic?

9 Optimization and Equality Constraints and Nonlinear Programming

In some maximization problems, the agents can choose only values of $(x_1, \ldots, x_n)$ that satisfy certain equalities. For example, a consumer has a budget constraint $p_1x_1 + \cdots + p_nx_n = m$.

$$\max_{x_1,\ldots,x_n} U(x) = U(x_1, \ldots, x_n) \text{ subject to } p_1x_1 + \cdots + p_nx_n = m.$$

The procedure is the same as before. (1) Use FOC to find critical points; (2) use SOC to check whether a critical point is a local maximum, or show that $U(X)$ is quasi-concave so that a critical point must be a global maximum.
Define $B = \{(x_1, \ldots, x_n) \text{ such that } p_1x_1 + \cdots + p_nx_n = m\}$.
A local maximum $X^* = (x_1^*, \ldots, x_n^*)$: $\exists \epsilon > 0$ such that $U(X^*) \geq U(X)$ for all $X \in B$ satisfying $x_i \in (x_i^* - \epsilon, x_i^* + \epsilon)\ \forall i$.
A critical point: A $X^* \in B$ satisfying the FOC for a local maximum.
A global maximum X^*: $F(X^*) \geq F(X)\ \ \forall X \in B$.
A unique global maximum X^*: $F(X^*) > F(X)\ \ \forall X \in B,\ X \neq X^*$.
To define the concept of a critical point, we have to know what is the FOC first.

9.1 FOC and SOC for a constraint maximization

Consider the 2-variable utility maximization problem:

$$\max_{x_1,x_2} U(x_1, x_2) \quad \text{subject to } p_1x_1 + p_2x_2 = m.$$

Using the budget constraint, $x_2 = \dfrac{m - p_1x_1}{p_2} = h(x_1)$, $\dfrac{dx_2}{dx_1} = -\dfrac{p_1}{p_2} = h'(x_1)$, and it becomes a single variable maximization problem:

$$\max_{x_1} U\left(x_1, \frac{m - p_1x_1}{p_2}\right), \quad \text{FOC: } \frac{dU}{dx_1} = U_1 + U_2\left(-\frac{p_1}{p_2}\right) = 0, \ \text{ SOC: } \frac{d^2U}{dx_1^2} < 0.$$

$$\frac{d^2U}{dx_1^2} = U_{11} - 2\frac{p_1}{p_2}U_{12} + \left(\frac{p_1}{p_2}\right)^2 U_{22} = \frac{-1}{p_2^2}\begin{vmatrix} 0 & -p_1 & -p_2 \\ -p_1 & U_{11} & U_{12} \\ -p_2 & U_{21} & U_{22} \end{vmatrix}.$$

By FOC, $\dfrac{U_1}{p_1} = \dfrac{U_2}{p_2} \equiv \lambda$ (MU of \$1).

$$\text{FOC: } U_1 = p_1\lambda,\ U_2 = p_2\lambda, \ \text{ SOC: } \begin{vmatrix} 0 & U_1 & U_2 \\ U_1 & U_{11} & U_{12} \\ U_2 & U_{21} & U_{22} \end{vmatrix} > 0.$$

Alternatively, we can define Lagrangian:

$$\mathcal{L} \equiv U(x_1, x_2) + \lambda(m - p_1x_1 - p_2x_2)$$

FOC: $\mathcal{L}_1 = \dfrac{\partial \mathcal{L}}{\partial x_1} = U_1 - \lambda p_1 = 0$, $\mathcal{L}_2 = \dfrac{\partial \mathcal{L}}{\partial x_2} = U_2 - \lambda p_2 = 0$, $\mathcal{L}_\lambda = \dfrac{\partial \mathcal{L}}{\partial \lambda} = m - p_1 x_1 - p_2 x_2 = 0$.

SOC: $\begin{vmatrix} 0 & -p_1 & -p_2 \\ -p_1 & \mathcal{L}_{11} & \mathcal{L}_{12} \\ -p_2 & \mathcal{L}_{21} & \mathcal{L}_{22} \end{vmatrix} > 0$.

general 2-variable with 1-constraint case:

$$\max_{x_1,x_2} F(x_1,x_2) \quad \text{subject to} \quad g(x_1,x_2) = 0.$$

Using the constraint, $x_2 = h(x_1)$, $\dfrac{dx_2}{dx_1} = -\dfrac{g_1}{g_2} = h'(x_1)$, and it becomes a single variable maximization problem:

$$\max_{x_1} F\left(x_1, h(x_1)\right) \quad \text{FOC: } \frac{dF}{dx_1} = F_1 + F_2 h'(x_1) = 0, \quad \text{SOC: } \frac{d^2F}{dx_1^2} < 0.$$

$$\frac{d^2F}{dx_1^2} = \frac{d}{dx_1}\left(F_1 + F_2 h'\right) = F_{11} + 2h'F_{12} + \left(h'\right)^2 F_{22} + F_2 h''$$

$$h'' = \frac{d}{dx_1}\left(-\frac{g_1}{g_2}\right) = \frac{-1}{g_2}\left[g_{11} + 2g_{12}h' + g_{22}(h')^2\right].$$

$$\Rightarrow \frac{d^2F}{dx_1^2} = \left(F_{11} - \frac{F_2}{g_2}g_{11}\right) + 2\left(F_{12} - \frac{F_2}{g_2}g_{12}\right)h' + \left(F_{22} - \frac{F_2}{g_2}g_{22}\right)(h')^2$$

$$= -\begin{vmatrix} 0 & -h' & 1 \\ -h' & F_{11} - \frac{F_2}{g_2}g_{11} & F_{12} - \frac{F_2}{g_2}g_{12} \\ 1 & F_{21} - \frac{F_2}{g_2}g_{21} & F_{22} - \frac{F_2}{g_2}g_{22} \end{vmatrix} = \frac{-1}{g_2^2}\begin{vmatrix} 0 & g_1 & g_2 \\ g_1 & F_{11} - \frac{F_2}{g_2}g_{11} & F_{12} - \frac{F_2}{g_2}g_{12} \\ g_2 & F_{21} - \frac{F_2}{g_2}g_{21} & F_{22} - \frac{F_2}{g_2}g_{22} \end{vmatrix}$$

By FOC, $\dfrac{F_1}{g_1} = \dfrac{F_2}{g_2} \equiv \lambda$ (Lagrange multiplier).

Alternatively, we can define Lagrangian:

$$\mathcal{L} \equiv F(x_1,x_2) - \lambda g(x_1,x_2)$$

FOC: $\mathcal{L}_1 = \dfrac{\partial \mathcal{L}}{\partial x_1} = F_1 - \lambda g_1 = 0$, $\mathcal{L}_2 = \dfrac{\partial \mathcal{L}}{\partial x_2} = F_2 - \lambda g_2 = 0$, $\mathcal{L}_\lambda = \dfrac{\partial \mathcal{L}}{\partial \lambda} = -g(x_1,x_2) = 0$.

SOC: $\begin{vmatrix} 0 & g_1 & g_2 \\ g_1 & \mathcal{L}_{11} & \mathcal{L}_{12} \\ g_2 & \mathcal{L}_{21} & \mathcal{L}_{22} \end{vmatrix} > 0$.

n-variable 1-constraint case:

$$\max_{x_1,\ldots,x_n} F(x_1,\ldots,x_n) \quad \text{subject to} \quad g(x_1,\ldots,x_n) = 0.$$

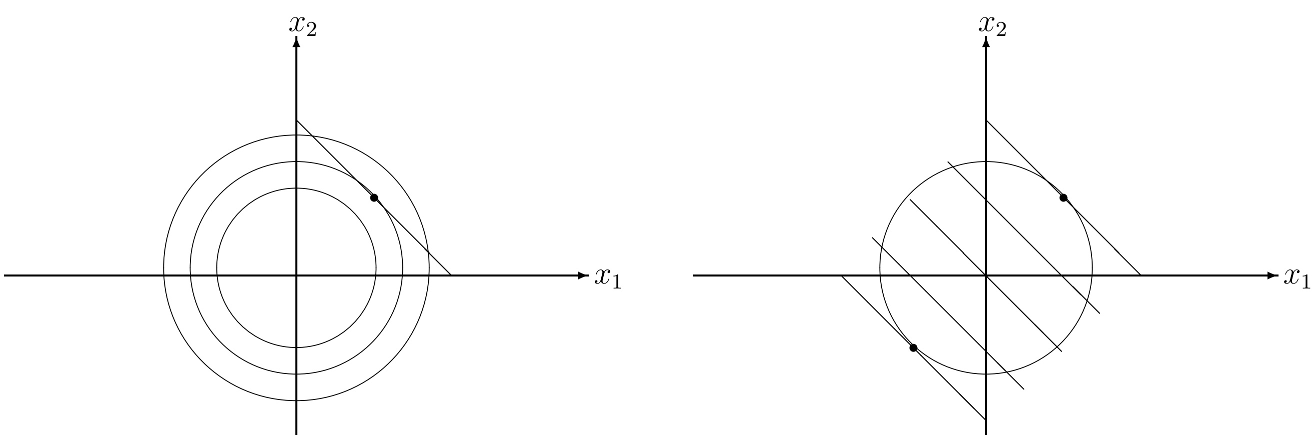
$$\mathcal{L} \equiv F(x_1,\ldots,x_n) - \lambda g(x_1,\ldots,x_n)$$

FOC: $\mathcal{L}_i = \dfrac{\partial \mathcal{L}}{\partial x_i} = F_i - \lambda g_i = 0$, $i = 1, \ldots, n$, $\mathcal{L}_\lambda = \dfrac{\partial \mathcal{L}}{\partial \lambda} = -g(x_1, \ldots, x_n) = 0$.

SOC: $\begin{vmatrix} 0 & -g_1 \\ -g_1 & \mathcal{L}_{11} \end{vmatrix} < 0$, $\begin{vmatrix} 0 & -g_1 & -g_2 \\ -g_1 & \mathcal{L}_{11} & \mathcal{L}_{12} \\ -g_2 & \mathcal{L}_{21} & \mathcal{L}_{22} \end{vmatrix} > 0$ $\begin{vmatrix} 0 & -g_1 & -g_2 & -g_3 \\ -g_1 & \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} \\ -g_2 & \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} \\ -g_3 & \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} \end{vmatrix} < 0$, etc.

9.2 Examples

Example 1: $\max F(x_1, x_2) = -x_1^2 - x_2^2$ subject to $x_1 + x_2 = 1$.
$\mathcal{L} = -x_1^2 - x_2^2 + \lambda(1 - x_1 - x_2)$.
FOC: $\mathcal{L}_1 = -2x_1 - \lambda = 0$, $\mathcal{L}_2 = -2x_2 - \lambda = 0$ and $\mathcal{L}_\lambda = 1 - 2x_1 - 2x_2 = 0$.
Critical point: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$, $\lambda = -1$.
SOC: $\begin{vmatrix} 0 & 1 & 1 \\ 1 & -2 & 0 \\ 1 & 0 & -2 \end{vmatrix} = 4 > 0$.
$\Rightarrow \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ is a local maximum.



Example 2: $\max F(x_1, x_2) = x_1 + x_2$ subject to $x_1^2 + x_2^2 = 1$.
$\mathcal{L} = x_1 + x_2 + \lambda(1 - x_1^2 - x_2^2)$.
FOC: $\mathcal{L}_1 = 1 - 2\lambda x_1 = 0$, $\mathcal{L}_2 = 1 - 2\lambda x_2 = 0$ and $\mathcal{L}_\lambda = 1 - x_1^2 - x_2^2 = 0$.
Two critical points: $x_1 = x_2 = \lambda = 1/\sqrt{2}$ and $x_1 = x_2 = \lambda = -1/\sqrt{2}$.
SOC: $\begin{vmatrix} 0 & -2x_1 & -2x_2 \\ -2x_1 & -2\lambda & 0 \\ -2x_2 & 0 & -2\lambda \end{vmatrix} = 8\lambda(x_1^2 + x_2^2)$.
$\Rightarrow x_1 = x_2 = \lambda = 1/\sqrt{2}$ is a local maximum and $x_1 = x_2 = \lambda = -1/\sqrt{2}$ is a local minimum.

9.3 Cost minimization and cost function

$\min C(x_1, x_2; w_1, w_2) = w_1x_1 + w_2x_2$ subject to $x_1^a x_2^{1-a} = q$, $0 < a < 1$.
$\mathcal{L} = w_1x_1 + w_2x_2 + \lambda(q - x_1^a x_2^{1-a})$.
FOC: $\mathcal{L}_1 = w_1 - a\lambda x_1^{a-1}x_2^{1-a} = 0$, $\mathcal{L}_2 = w_2 - (1-a)\lambda x_1^a x_2^{-a} = 0$, $\mathcal{L}_\lambda = q - x_1^a x_2^{1-a} = 0$.
$\Rightarrow \dfrac{w_1}{w_2} = \dfrac{ax_2}{(1-a)x_1} \Rightarrow x_1 = q\left[\dfrac{aw_2}{(1-a)w_1}\right]^{1-a}$, $x_2 = q\left[\dfrac{(1-a)w_1}{aw_2}\right]^a$.
SOC:

$$\begin{vmatrix} 0 & -ax_1^{a-1}x_2^{1-a} & -(1-a)x_1^a x_2^{-a} \\ -ax_1^{a-1}x_2^{1-a} & a(1-a)\lambda x_1^{a-2}x_2^{1-a} & -a(1-a)\lambda x_1^{a-1}x_2^{-a} \\ -(1-a)x_1^a x_2^{-a} & -a(1-a)\lambda x_1^{a-1}x_2^{-a} & a(1-a)\lambda x_1^a x_2^{-a-1} \end{vmatrix} = -\frac{a(1-a)q^3\lambda}{(x_1x_2)^2} < 0.$$

$\Rightarrow x_1 = q\left[\dfrac{aw_2}{(1-a)w_1}\right]^{1-a}$, $x_2 = q\left[\dfrac{(1-a)w_1}{aw_2}\right]^a$ is a local minimum.
The total cost is $C(w_1, w_2, q) = q\left[\left(\dfrac{a}{1-a}\right)^{1-a} + \left(\dfrac{1-a}{a}\right)^a\right] w_1^a w_2^{1-a}$.

9.4 Utility maximization and demand function

$\max U(x_1, x_2) = a\ln x_1 + b\ln x_2$ subject to $p_1x_1 + p_2x_2 = m$.
$\mathcal{L} = a\ln x_1 + b\ln x_2 + \lambda(m - p_1x_1 - p_2x_2)$.
FOC: $\mathcal{L}_1 = \dfrac{a}{x_1} - \lambda p_1 = 0$, $\mathcal{L}_2 = \dfrac{b}{x_2} - \lambda p_2 = 0$ and $\mathcal{L}_\lambda = m - p_1x_1 - p_2x_2 = 0$.
$\Rightarrow \dfrac{a}{b}\dfrac{x_2}{x_1} = \dfrac{p_1}{p_2} \Rightarrow x_1 = \dfrac{am}{(a+b)p_1}$, $x_2 = \dfrac{bm}{(a+b)p_2}$

$$\text{SOC: } \begin{vmatrix} 0 & -p_1 & -p_2 \\ -p_1 & \dfrac{-a}{x_1^2} & 0 \\ -p_2 & 0 & \dfrac{-b}{x_2^2} \end{vmatrix} = \frac{ap_2^2}{x_1^2} + \frac{bp_1^2}{x_2^2} > 0.$$

$\Rightarrow x_1 = \dfrac{am}{(a+b)p_1}$, $x_2 = \dfrac{bm}{(a+b)p_2}$ is a local maximum.

9.5 Quasi-concavity and quasi-convexity

As discussed in the intermediate microeconomics course, if the MRS is strictly decreasing along an indifference curve (indifference curve is convex toward the origin), then the utility maximization has a unique solution. A utility function $U(x_1, x_2)$ is **quasi-concave** if MRS $(= \dfrac{U_1}{U_2})$ is decreasing along every indifference curve. In case MRS is strictly decreasing, the utility function is **strictly quasi-concave**. If $U(x_1, x_2)$ is (strictly) quasi-concave, then a critical point must be a (unique) global maximum.

Two ways to determine whether $U(x_1, x_2)$ is quasi-concave: (1) the set $\{(x_1, x_2) \in R^2\ U(x_1, x_2) \geq \bar{U}\}$ is convex for all $\bar{U}$. (Every indifference curve is convex toward the origin.)

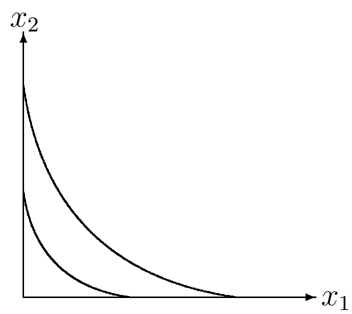
(2) $\displaystyle \frac{d}{dx_1}\left(-\frac{U_1}{U_2}\right)\Bigg|_{U(x_1,x_2)=\bar{U}} > 0$. (MRS is strictly decreasing along every indifference curve.)

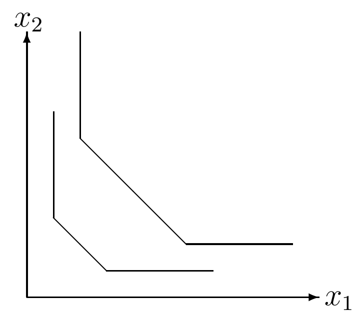
In sections 6.14 and 7.7 we used $F(X)$ to define two sets, G_F^+ and G_F^-, and we say that $F(X)$ is concave (convex) if G_F^- (G_F^+) is a convex set. Now we say that $F(X)$ is quasi-concave (quasi-convex) if every y-cross-section of G_F^- (G_F^+) is a convex set. A y-cross-section is formally defined as

$$G_F^-(\bar{y}) \equiv \{(x_1,x_2)|\ F(x_1,x_2) \ge \bar{y}\} \subset R^2, \quad G_F^+(\bar{y}) \equiv \{(x_1,x_2)|\ F(x_1,x_2) \le \bar{y}\} \subset R^2.$$

Clearly, G_F^- is the union of all $G_F^-(\bar{y})$: $G_F^- = \cup_{\bar{y}} G_F^-(\bar{y})$.

Formal definition: $F(x_1,\ldots,x_n)$ is **quasi-concave** if $\forall X^0, X^1 \in A$ and $0 \le \theta \le 1$, $F(X^\theta) \ge \min\{F(X^0), F(X^1)\}$.
$F(x_1,\ldots,x_n)$ is **strictly quasi-concave** if $\forall X^0 \ne X^1 \in A$ and $0 < \theta < 1$, $F(X^\theta) > \min\{F(X^0), F(X^1)\}$.





If F is (strictly) concave, then F must be (strictly) quasi-concave.
Proof: If $F(X^\theta) < \min\{F(X^0), F(X^1)\}$, then $F(X^\theta) < (1-\theta)F(X^0) + \theta F(X^1)$.
Geometrically, if G_F^- is convex, then $G_F^-(\bar{y})$ must be convex for every $\bar{y}$.
If F is quasi-concave, it is not necessarily that F is concave.
Counterexample: $F(x_1x_2) = x_1x_2,\ x_1,x_2 > 0$ is strictly quasi-concave but not concave. In this case, every $G_F^-(\bar{y})$ is convex but still G_F^- is not convex.

Bordered Hessian: $|B^1| \equiv \begin{vmatrix} 0 & F_1 \\ F_1 & F_{11} \end{vmatrix}$, $|B^2| \equiv \begin{vmatrix} 0 & F_1 & F_2 \\ F_1 & F_{11} & F_{12} \\ F_2 & F_{21} & F_{22} \end{vmatrix}$, $|B^3| \equiv \begin{vmatrix} 0 & F_1 & F_2 & F_3 \\ F_1 & F_{11} & F_{12} & F_{13} \\ F_2 & F_{21} & F_{22} & F_{23} \\ F_3 & F_{31} & F_{32} & F_{33} \end{vmatrix}$, etc.

Theorem 1: Suppose that F is twice differentiable. If F is quasi-concave, then $|B^2| \ge 0$, $|B^3| \le 0$, etc. Conversely, if $|B^1| < 0$, $|B^2| > 0$, $|B^3| < 0$, etc., then F is strictly quasi-concave.
Proof ($n=2$): $\displaystyle \frac{d\text{MRS}}{dx_1} = \frac{|B^2|}{F_2^3}$ and therefore F is quasi-concave if and only if $|B^2| \ge 0$.

Consider the following maximization problem with a **linear** constraint.

$$\max_{x_1,\ldots,x_n} F(x_1,\ldots,x_n) \quad \text{subject to } \; a1x_1+\cdots+a_nx_n=b.$$

Theorem 2: If F is (strictly) quasi-concave and X^0 satisfies FOC, then X^0 is a (unique) global maximum.
Proof: By theorem 1, X^0 must be a local maximum. Suppose there exists X^1 satisfying the linear constraint with $U(X^1) > U(X^0)$. Then $U(X^\theta) > U(X^0)$, a contradiction.

Theorem 3: A monotonic increasing transformation of a quasi-concave function is a quasi-concave function. A quasi-concave function is a monotonic increasing transformation of a concave function.
Proof: A monotonic increasing transformation does not change the sets $\{(x_1,x_2)\in R^2\; U(x_1,x_2)\ge \bar{U}\}$. To show the opposite, suppose that $f(x_1,x_2)$ is quasi-concave. Define a monotonic transformation as

$$g(x_1,x_2)=H(f(x_1,x_2)) \quad \text{where } H^{-1}(g)=f(x,x).$$

9.6 Elasticity of Substitution

Consider a production function $Q=F(x_1,x_2)$.
Cost minimization $\Rightarrow \dfrac{w_1}{w_2}=\dfrac{F_1}{F_2}\equiv\theta$. Let $\dfrac{x_2}{x_1}\equiv r$. On an isoquant $F(x_1,x_2)=\bar{Q}$, there is a relationship $r=\phi(\theta)$. The elasticity of substitution is defined as $\sigma\equiv\dfrac{\theta}{r}\dfrac{dr}{d\theta}$.
If the input price ratio $\theta=\dfrac{w_1}{w_2}$ increases by 1%, a competitive producer will increase its input ratio $r=\dfrac{x_2}{x_1}$ by σ%.
Example: For a CES function $Q=\{ax_1^\rho+bx_2^\rho\}^{\frac{k}{\rho}}$, $\sigma=\dfrac{1}{1-\rho}$.

9.7 Problems

1. Let $A=\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$, $B=\begin{pmatrix} 0 & c_1 & c_2\\ c_1 & a_{11} & a_{12}\\ c_2 & a_{21} & a_{22}\end{pmatrix}$, and $C=\begin{pmatrix} -c_2\\ c_1\end{pmatrix}$.
   (a) Calculate the determinant $|B|$.
   (b) Calculate the product $C'AC$. Does $|B|$ have any relationship with $C'AC$?
2. Consider the utility function $U(x_1,x_2)=x_1^\alpha x_2^\beta$.
   (a) Calculate the marginal utilities $U_1=\dfrac{\partial U}{\partial x_1}$ and $U_2=\dfrac{\partial U}{\partial x_2}$.
   (b) Calculate the hessian matrix $H=\begin{pmatrix} U_{11} & U_{12}\\ U_{21} & U_{22}\end{pmatrix}$.

(c) Calculate the determinant of the matrix $B = \begin{pmatrix} 0 & U_1 & U_2 \\ U_1 & U_{11} & U_{12} \\ U_2 & U_{21} & U_{22} \end{pmatrix}$.

(d) Let $C = \begin{pmatrix} -U_2 \\ U_1 \end{pmatrix}$. Calculate the product C'HC.

3. Use the Lagrangian multiplier method to find the critical points of $f(x, y) = x + y$ subject to $x^2 + y^2 = 2$. Then use the bordered Hessian to determine which point is a maximum and which is a minimum. (There are two critical points.)

4. Determine whether $f(x, y) = xy$ is concave, convex, quasiconcave, or quasiconvex. ($x > 0$ and $y > 0$.)

5. Let $U(x, y)$, $U, x, y > 0$, be a homogenous of degree 1 and concave function with $U_{xy} \neq 0$.

   (a) Show that $V(x, y) = [U(x, y)]^a$ is strictly concave if $0 < a < 1$ and $V(x, y)$ is strictly quasi-concave for all $a > 0$. Hint: $U_{xx}U_{yy} = [U_{xy}]^2$.

   (b) Show that $F(x, y) = \left(x^\beta + y^\beta\right)^{a/\beta}$, $x, y > 0$ and $-\infty < \beta < 1$ is homogenous of degree 1 and concave if $a = 1$.

   (c) Determine the range of a so that $F(x, y)$ is strictly concave.

   (d) Determine the range of a so that $F(x, y)$ is strictly quasi-concave.

9.8 Nonlinear Programming

The general nonlinear programming problem is:

$$\max_{x_1,\ldots,x_n} F(x_1,\ldots,x_n) \quad \text{subject to} \quad \begin{cases} g^1(x_1,\ldots,x_n) \le b_1 \\ \vdots \\ g^m(x_1,\ldots,x_n) \le b_m \\ x_1,\ldots,x_n \ge 0. \end{cases}$$

In equality constraint problems, the number of constraints should be less than the number of policy variables, $m < n$. For nonlinear programming problems, there is no such a restriction, m can be greater than or equal to n. In vector notation,

$$\max_x F(x) \quad \text{subject to} \quad g(x) \le b, \quad x \ge 0.$$

9.9 Kuhn-Tucker condition

Define the Lagrangian function as

$$L(x,y) = F(x) + y(b - g(x)) = F(x_1,\ldots,x_n) + \sum_{j=1}^{m} y_j(b_j - g^j(x_1,\ldots,x_n)).$$

Kuhn-Tucker condition: The FOC is given by the Kuhn-Tucker conditions:

$$\frac{\partial L}{\partial x_i} = \frac{\partial F}{\partial x_i} - \sum_{j=1}^{m} y_j \frac{\partial g^j}{\partial x_i} \le 0, \quad x_i \frac{\partial L}{\partial x_i} = 0 \quad x_i \ge 0, \quad i = 1,\ldots,n$$

$$\frac{\partial L}{\partial y_j} = b_j - g^j(x) \ge 0, \quad y_j \frac{\partial L}{\partial y_j} = 0 \quad y_j \ge 0, \quad j = 1,\ldots,m$$

Kuhn Tucker theorem: x^* solves the nonlinear programming problem if (x^*, y^*) solves the saddle point problem:

$$L(x, y^*) \le L(x^*, y^*) \le L(x^*, y) \quad \text{for all} \quad x \ge 0, \quad y \ge 0,$$

Conversely, suppose that $f(x)$ is a concave function and $g^j(x)$ are convex functions (concave programming) and the constraints satisfy the constraint qualification condition that there is some point in the opportunity set which satisfies all the inequality constraints as strict inequalities, i.e., there exists a vector $x^0 \ge 0$ such that $g^j(x^0) < b_j$, $j = 1,\ldots,m$, then x^* solves the nonlinear programming problem only if there is a y^* for which (x^*, y^*) solves the saddle point problem.

If constraint qualification is not satisfied, it is possible that a solution does not satisfy the K-T condition. If it is satisfied, then the K-T condition will be necessary. For the case of concave programming, it is also sufficient.

In economics applications, however, it is not convenient to use K-T condition to find the solution. In stead, we first solve the equality constraint version of the problem and then use K-T condition to check or modify the solution when some constraints are violated.

The K-T condition for minimization problems: the inequalities reversed.

9.10 Examples

Example 1. (Joint product profit maximization) The cost function of a competitive producer producing 2 joint products is $c(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$. The profit function is given by $\pi(p_1, p_2) = p_1x_1 + p_2x_2 - (x_1^2 + x_1x_2 + x_2^2)$.

$$\max_{x_1 \geq 0, x_2 \geq 0} f(x_1, x_2) = p_1x_1 + p_2x_2 - (x_1^2 + x_1x_2 - x_2^2)$$

K-T condition: $f_1 = p_1 - 2x_1 - x_2 \leq 0$, $f_2 = p_2 - x_1 - 2x_2 \leq 0$, $x_if_i = 0$, $i = 1, 2$.
Case 1. $p_1/2 < p_2 < 2p_1$.
$x_1 = (2p_1 - p_2)/3$, $x_2 = (2p_2 - p_1)/3$.

Case 2. $2p_1 < p_2$.
$x_1 = 0$, $x_2 = p_2/2$.

Case 3. $2p_2 < p_1$.
$x_1 = p_1/2$, $x_2 = 0$.

Example 2. The production function of a producer is given by $q = (x_1+1)(x_2+1)-1$. For $q = 8$, calculate the cost function $c(w_1, w_2)$.

$$\min_{x_1 \geq 0, x_2 \geq 0} w_1x_1 + w_2x_2 \qquad \text{subject to} \qquad -[(x_1+1)(x_2+1) - 1] \geq -8$$

Lagrangian function: $L = w_1x_1 + w_1x_2 + \lambda[(x_1 + 1)(x_2 + 1) - 9]$.
K-T conditions: $L_1 = w_1 - \lambda(x_2 - 1) \geq 0$, $L_2 = w_2 - \lambda(x_2 - 1) \geq 0$, $x_iL_i = 0$, $i = 1, 2$, and $L_\lambda = (x_1 + 1)(x_2 + 1) - 9 \geq 0$, $\lambda L_\lambda = 0$.

Case 1. $w_1/9 < w_2 < 9w_1$.
$x_1 = \sqrt{9w_2/w_1} - 1$, $x_2 = \sqrt{9w_1/w_2} - 1$ and $c(w_1, w_2) = 6\sqrt{w_1w_2} - w_1 - w_2$.

Case 2. $9w_1 < w_2$.
$x_1 = 8$, $x_2 = 0$, and $c(w_1, w_2) = 8w_1$.

Case 3. $9w_2 < w_1$.
$x_1 = 0$, $x_2 = 8$, $c(w_1, w_2) = 8w_2$.

Example 3. The utility function of a consumer is $U(x_1, x_2) = x_1(x_2+1)$. The market price is $p_1 = p_2 = 1$ and the consumer has \$11. Therefore the budget constraint is $x_1 + x_2 \leq 11$. Suppose that both products are under rationing. Besides the money price, the consumer has to pay ρ_i rationing points for each unit of product i consumed. Assume that $\rho_1 = 1$ and $\rho_2 = 2$ and the consumer has q rationing points. The rationing point constraint is $x_1 + 2x_2 \leq q$. The utility maximization problem is given by

$$\max_{x_1,x_2} U(x_1, x_2) = x_1(x_2+1) \quad \text{subject to} \quad x_1+x_2 \leq 11, \quad x_1+2x_2 \leq q, \quad x_1, x_2 \geq 0.$$

Lagrangian function: $L = x_1(x_2 + 1) + \lambda_1(11 - x_1 - x_2) + \lambda_2(q - x_1 - 2x_2)$.
K-T conditions: $L_1 = x_2+1-\lambda_1-\lambda_2 \leq 0$, $L_2 = x_1-\lambda_1-2\lambda_2 \leq 0$, $x_iL_i = 0$, $i = 1, 2$, and $L_{\lambda_1} = 11 - x_1 - x_2 \geq 0$, $L_{\lambda_2} = q - x_1 - 2x_2 \geq 0$ $\lambda_i L_{\lambda_i} = 0$.

Case 1: $q < 2$.
$x_1 = q$, $x_2 = 0$, $\lambda_1 = 0$, and $\lambda_2 = 1$.

Case 2: $2 \leq q \leq 14$.
$x_1 = (q + 2)/2$, $x_2 = (q - 2)/4$, $\lambda_1 = 0$, and $\lambda_2 = (q + 2)/4$.

Case 3: $14 < q \leq 16$.
$x_1 = 22 - q$, $x_2 = q - 11$, $\lambda_1 = 3(q - 14)$, and $\lambda_2 = 2(16 - q)$.

Case 4: $16 < q$.
$x_1 = 6$, $x_2 = 5$, $\lambda_1 = 6$, and $\lambda_2 = 0$.

9.11 Problems

1. Given the individual utility function $U(X,Y) = 2\sqrt{X} + Y$,
a) show that U is quasi-concave for $X \geq 0$ and $Y \geq 0$,
b) state the Kuhn-Tucker conditions of the following problem:

$$\max_{X\geq 0,\, Y\geq 0} 2\sqrt{X} + Y$$

$$\text{s. t. } \; P_X X + P_Y Y \leq I,$$

c) derive the demand functions $X(P_X, P_Y, I)$ and $Y(P_X, P_Y, I)$ for the case $I \geq \frac{P_Y^2}{P_X}$, check that the K-T conditions are satisfied,
d) and do the same for $I < \frac{P_Y^2}{P_X}$.
e) Given that $I = 1$ and $P_Y = 1$, derive the ordinary demand function $X = D(P_X)$.
f) Are your answers in (c) and (d) global maximum? Unique global maximum? Why

or why not?

2. A farm has a total amount of agricultural land of one acre. It can produce two crops, corn (C) and lettuce (L), according to the production functions $C = N_C$ and $L = 2\sqrt{N_L}$ respectively, where N_C (N_L) is land used in corn (lettuce) production. The prices of corn and lettuce are p and q respectively. Thus, if the farm uses N_C of land in corn production and N_L in lettuce production, ($N_C \geq 0$, $N_L \geq 0$, and $N_C + N_L \leq 1$) its total revenue is $pN_C + 2q\sqrt{N_L}$.
a) Suppose the farm is interested in maximizing its revenue. State the revenue maximization problem and the Kuhn-Tucker conditions.
b) Given that $q > p > 0$, how much of each output will the farm produce? Check that the K-T conditions are satisfied.
c) Given that $p \geq q > 0$, do the same as (b).

3. Suppose that a firm has two activities producing two goods "meat" (M) and "egg" (E) from the same input "chicken" (C) according to the production functions $M = C_M$ and $E = C_E^{1/2}$, where C_M (respectively C_E) ≥ 0 is the q. Suppose in the short run, the firm has $\bar{C}$ units of chicken that it must take as given and suppose that the firm faces prices $p > 0$, $q > 0$ of meat and egg respectively.
a) Show that the profit function $\pi = pC_M + qC_E^{0.5}$ is quasi-concave in (C_M, C_E).
b) Write down the short run profit maximization problem.
c) State the Kuhn-Tucker conditions.
d) Derive the short run supply functions. (There are two cases.)
e) Is your solution a global maximum? Explain.

4. State the Kuhn-Tucker conditions of the following nonlinear programming problem

$$\begin{aligned} \max \quad & U(X,Y) = 3\ln X + \ln Y \\ \text{s. t.} \quad & 2X + Y \leq 24 \\ & X + 2Y \leq 24 \\ & X \geq 0, \quad Y \geq 0. \end{aligned}$$

Show that $X = 9$, $Y = 6$, $\lambda_1 = 1/6$, and $\lambda_2 = 0$ satisfy the Kuhn-Tucker conditions. What is the economic interpretations of $\lambda_1 = 1/6$ and $\lambda_2 = 0$ if the first constraint is interpreted as the income constraint and the second constraint as the rationing point constraint of a utility maximization problem?

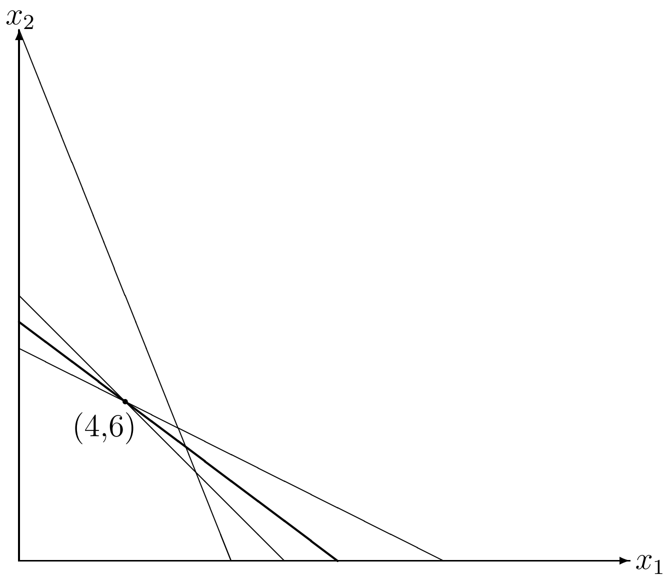
9.12 Linear Programming – A Graphic Approach

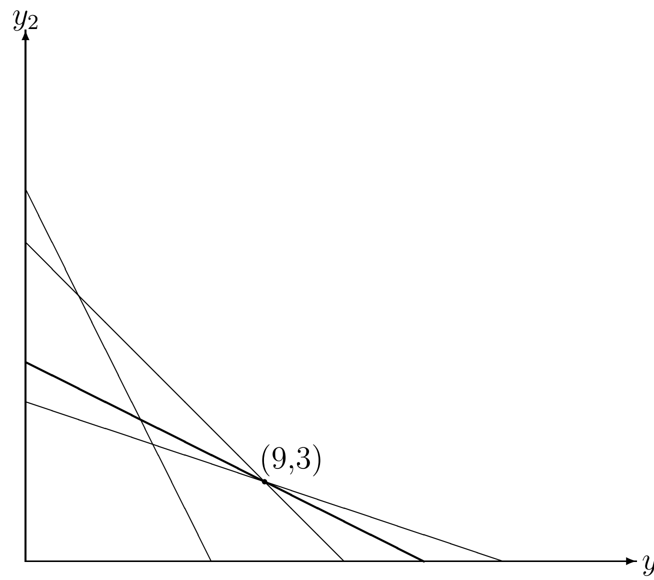
Example 1 (A Production Problem). A manufacturer produces tables x_1 and desks x_2. Each table requires 2.5 hours for assembling (A), 3 hours for buffing (B), and 1 hour for crating (C). Each desks requires 1 hour for assembling (A), 3 hours for buffing (B), and 2 hours for crating (C). The firm can use no more than 20 hours for assembling, 30 hours for buffing, and 16 hours for crating each week. Its profit margin is \$3 per table and \$4 per desk.

$$\begin{aligned} \max \Pi &= 3x_1 + 4x_2 && (4) \\ \text{subject to} \quad 2.5x_1 + x_2 &\leq 20 && (5) \\ 3x_1 + 3x_2 &\leq 30 && (6) \\ x_1 + 2x_2 &\leq 16 && (7) \\ x_1, x_2 &\geq 0. && (8) \end{aligned}$$

extreme point: The intersection of two constraints.
extreme point theorem: If an optimal feasible value of the objective function exists, it will be found at one of the extreme points.

In the example, There are 10 extreme points, but only 5 are feasible: (0 ,0), (8, 0), ($6\frac{2}{3}$, $3\frac{1}{3}$), (4, 6), and (0, 8), called basic feasible solutions. At (4, 6), $\Pi = 36$ is the optimal.





Example 2 (The Diet Problem). A farmer wants to see that her herd gets the minimum daily requirement of three basic nutrients A, B, and C. Daily requirements are 14 for A, 12 for B, and 18 for C. Product y_1 has 2 units of A and 1 unit each of B and C; product y_2 has 1 unit each of A and B and 3 units of C. The cost of y_1 is \$2, and the cost of y_2 is \$4.

$$\begin{aligned} \min c &= 2y_1 + 4y_2 && (9) \\ \text{subject to} \quad 2y_1 + y_2 &\geq 14 && (10) \\ y_1 + y_2 &\geq 12 && (11) \\ y_1 + 3y_2 &\geq 18 && (12) \\ y_1, y_2 &\geq 0. && (13) \end{aligned}$$

Slack and surplus variables: To find basic solutions, equations are needed. This is done by incorporating a separate slack or surplus variable s_i into each inequality.
In example 1, the system becomes

$$2.5x_1 + x_2 + s_1 = 20 \qquad 3x_1 + 3x_2 + s_2 = 30 \qquad x_1 + 2x_2 + s_3 = 16.$$

In matrix for,

$$\begin{bmatrix} 2.5 & 1 & 1 & 0 & 0 \\ 3 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 30 \\ 16 \end{bmatrix}.$$

In example 2, the inequalities are "$\geq$" and the surplus variables are substracted:

$$2y_1 + y_2 - s_1 = 14 \qquad y_1 + y_2 - s_2 = 12 \qquad y_1 + 3y_2 - s_3 = 18.$$

In matrix for,

$$\begin{bmatrix} 2 & 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 \\ 1 & 3 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 12 \\ 18 \end{bmatrix}.$$

For a system of m equations and n variables, where $n > m$, a solution in which at least $n - m$ variables equal to zero is an extreme point. Thus by setting $n - m$and solving the m equations for the remaining m variables, an extreme point can be found. There are $n!/m!(n-m)!$ such solutions.

9.13 Linear programming – The simplex algorithm

The algorithm moves from one basic feasible solution to another, always improving upon the previous solutions, until the optimal solution is reached. In each step, those variables set equal to zero are called not in the basis and those not set equal to zero are called in the basis. Let use example one to illustrate the procedure.

1. The initial Simplex Tableau

x_1	x_2	s_1	s_2	s_3	Constant
2.5	1	1	0	0	20
3	3	0	1	0	30
1	2	0	0	1	16
-3	-4	0	0	0	0

   The first basic feasible solution can be read from the tableau as $x_1 = 0$, $x_2 = 0$, $s_1 = 20$, $s_2 = 30$, and $s_3 = 16$. The value of Π is zero.

2. The Pivot Element and a change of Basis

(a) The negative indicator with the largest absolute value determines the variable to enter the basis. Here it is x_2. The x_2 column is called the pivot column (j-th column).

(b) The variable to be eliminated is determined by the smallest displacement ratio. Displacement ratios are found by dividing the elements of the constant column by the elements of the pivot column. Here the smallest is 16/2=8 and row 3 is the pivot row (i-th row). The pivot element is 2.

3. (Pivoting) and we are going to move to the new basic solution with $s_3 = 0$ and $x_2 > 0$. First, divides every element of the pivoting row by the pivoting element (2 in this example) to make the pivoting element equal to 1. Then subtracts a_{kj} times the pivoting row from k-th row to make the j-th column a unit vector. (This procedure is called the Gaussian elimination method, usually used in solving simultaneous equations.) After pivoting, the Tableau becomes

x_1	x_2	s_1	s_2	s_3	Constant
2	0	1	0	-.5	12
1.5	0	0	1	-1.5	6
.5	1	0	0	.5	8
-1	0	0	0	2	32

The basic solution is $x_2 = 8$, $s_1 = 12$, $s_2 = 6$, and $x_1 = s_3 = 0$. The value of Π is 32.

4. (Optimization) Repeat steps 2-3 until a maximum is reached. In the example, x_1 column is the new pivoting column. The second row is the pivoting row and 1.5 is the pivoting element. After pivoting, the tableau becomes

x_1	x_2	s_1	s_2	s_3	Constant
0	0	1	$-\frac{4}{3}$	1.5	4
1	0	0	$\frac{2}{3}$	-1	4
0	1	0	$-\frac{1}{3}$	1	6
0	0	0	$\frac{2}{3}$	1	36

The basic solution is $x_1 = 4$, $x_2 = 6$, $s_1 = 4$, and $s_2 = s_3 = 0$. The value of Π is 36. Since there is no more negative indicators, the process stops and the basic solution is the optimal. $s_1 = 4 > 0$ and $s_2 = s_3 = 0$ means that the first constraint is not binding but the second and third are binding. The indicators for s_2 and s_3, $t_2 = \frac{2}{3}$ and $t_3 = 1$ are called the shadow values, representing the marginal contributions of increasing one hour for buffing or crating.

Because $y_1 = y_2 = 0$ is not feasible, the simplex algorithm for minimization problem is more complex. Usually, we solve its dual problem.

9.14 Linear programming – The dual problem

To every linear programming problem there corresponds a dual problem. If the original problem, called the primal problem, is

$$\max_x F = cx \quad \text{subject to} \quad Ax \le b, \quad x \ge 0$$

then the dual problem is

$$\min_y G = yb \quad \text{subject to} \quad yA \ge c, \quad y \ge 0$$

where

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad c = (c_1, \dots, c_n), \quad y = (y_1, \dots, y_m).$$

Existence theorem: A necessary and sufficient condition for the existence of a solution is that the opportunity sets of both the primal and dual problems are nonempty.
Proof: Suppose x, y are feasible. Then $Ax \le b$, $yA \ge c$. It follows that $F(x) = cx \le yAx$ and $G(y) = yb \ge yAx$. Therefore, $F(x) \le G(y)$.

Duality theorem: A necessary and sufficient condition for a feasible vector to represent a solution is that there exists a feasible vector for the dual problem for which the values of the objective functions of both problems are equal.

Complementary slackness theorem: A necessary and sufficient condition for feasible vectors x^* , y^* to solve the dual problems is that they satisfy the complementary slackness condition:

$$(c - y^*A)x^* = 0 \qquad y^*(b - Ax^*) = 0.$$

Proof: Use Kuhn-Tucker theorem.

Dual of the Diet Problem

$$\begin{aligned} \max c^* &= 14x_1 + 12x_2 + 18x_3 && (14) \\ \text{subject to} \quad 2x_1 + x_2 + x_1 &\le 2 && (15) \\ x_1 + x_2 + 3x_3 &\le 4 && (16) \\ x_1, x_2, x_3 &\ge 0. && (17) \end{aligned}$$

x_1, x_2, x_3, is interpreted as the imputed value of nutrient A, B, C, respectively.

10 General Equilibrium and Game Theory

10.1 Utility maximization and demand function

A consumer wants to maximize his utility function subject to his budget constraint:

$$\max U(x_1, \ldots, x_n) \qquad \text{subj. to } p_1x_1 + \cdots + p_nx_n = I.$$

Endogenous variables: $x_1, \ldots, x_n$
Exogenous variables: $p_1, \ldots, p_n, I$ (the consumer is a price taker)
Solution is the demand functions $x_k = D_k(p_1, \ldots, p_n, I)$, $k = 1, \ldots, n$

Example: $\max U(x_1, x_2) = a \ln x_1 + b \ln x_2$ subject to $p_1x_1 + p_2x_2 = m$.
$\mathcal{L} = a \ln x_1 + b \ln x_2 + \lambda(m - p_1x_1 - p_2x_2)$.
FOC: $\mathcal{L}_1 = \dfrac{a}{x_1} - \lambda p_1 = 0$, $\mathcal{L}_2 = \dfrac{b}{x_2} - \lambda p_2 = 0$ and $\mathcal{L}_\lambda = m - p_1x_1 - p_2x_2 = 0$.
$\Rightarrow \dfrac{a}{b}\dfrac{x_2}{x_1} = \dfrac{p_1}{p_2} \Rightarrow x_1 = \dfrac{am}{(a+b)p_1}$, $x_2 = \dfrac{bm}{(a+b)p_2}$

SOC: $\begin{vmatrix} 0 & -p_1 & -p_2 \\ -p_1 & \dfrac{-a}{x_1^2} & 0 \\ -p_2 & 0 & \dfrac{-b}{x_2^2} \end{vmatrix} = \dfrac{ap_2^2}{x_1^2} + \dfrac{bp_1^2}{x_2^2} > 0.$

$\Rightarrow x_1 = \dfrac{am}{(a+b)p_1}$, $x_2 = \dfrac{bm}{(a+b)p_2}$ is a local maximum.

10.2 Profit maximization and supply function

A producer's production technology can be represented by a production function $q = f(x_1, \ldots, x_n)$. Given the prices, the producer maximizes his profits:

$$\max \Pi(x_1, \ldots, x_n; p, p_1, \ldots, p_n) = pf(x_1, \ldots, x_n) - p_1x_1 - \cdots - p_nx_n$$

Exogenous variables: $p, p_1, \ldots, p_n$ (the producer is a price taker)
Solution is the supply function $q = S(p, p_1, \ldots, p_n)$ and the input demand functions, $x_k = X_k(p, p_1, \ldots, p_n)$ $k = 1, \ldots, n$

Example: $q = f(x_1, x_2) = 2\sqrt{x_1} + 2\sqrt{x_2}$ and $\Pi(x_1, x_2; p, p_1, p_2) = p(2\sqrt{x_1} + 2\sqrt{x_2}) - p_1x_1 - p_2x_2$,

$$\max_{x_1 . x_2} p(2\sqrt{x_1} + 2\sqrt{x_2}) - p_1x_1 - p_2x_2$$

FOC: $\dfrac{\partial \Pi}{\partial x_1} = \dfrac{p}{\sqrt{x_1}} - p_1 = 0$ and $\dfrac{\partial \Pi}{\partial x_2} = \dfrac{p}{\sqrt{x_2}} - p_2 = 0$.
$\Rightarrow x_1 = (p/p_1)^2$, $x_2 = (p/p_2)^2$ (input demand functions) and
$q = 2(p/p_1) + 2(p/p_2) = 2p(\frac{1}{p_1} + \frac{1}{p_2})$ (the supply function)

$\Pi = p^2(\frac{1}{p_1} + \frac{1}{p_2})$
SOC:

$$\begin{bmatrix} \frac{\partial^2 \Pi}{\partial x_1^2} & \frac{\partial^2 \Pi}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \Pi}{\partial x_1 \partial x_2} & \frac{\partial^2 \Pi}{\partial x_1^2} \end{bmatrix} = \begin{bmatrix} \frac{-p}{2x_1^{3/2}} & 0 \\ 0 & \frac{-p}{2x_2^{3/2}} \end{bmatrix}$$

is negative definite.

10.3 Transformation function and profit maximization

In more general cases, the technology of a producer is represented by a transformation function: $F^j(y_1^j, \ldots, y_n^j) = 0$, where $(y_1^j, \ldots, y_n^j)$ is called a production plan, if $y_k^j > 0$ (y_k^j) then k is an output (input) of j.

Example: a producer produces two outputs, y_1 and y_2, using one input y_3. Its technology is given by the transformation function $(y_1)^2 + (y_2)^2 + y_3 = 0$. Its profit is $\Pi = p_1 y_1 + p_2 y_2 + p_3 y_3$. The maximization problem is

$$\max_{y_1, y_2, y_3} p_1 y_1 + p_2 y_2 + p_3 y_3 \qquad \text{subject to } (y_1)^2 + (y_2)^2 + y_3 = 0.$$

To solve the maximization problem, we can eliminate y_3: $x = -y_3 = (y_1)^2 + (y_2)^2 > 0$ and

$$\max_{y_1, y_2} p_1 y_1 + p_2 y_2 - p_3[(y_1)^2 + (y_2)^2].$$

The solution is: $y_1 = p_1/(2p_3)$, $y_2 = p_2/(2p_3)$ (the supply functions of y_1 and y_2), and $x = -y_3 = [p_1/(2p_3)]^2 + [p_2/(2p_3)]^2$ (the input demand function for y_3).

10.4 The concept of an abstract economy and a competitive equilibrium

Commodity space: Assume that there are n commodities. The commodity space is $R_+^n = \{(x_1, \ldots, x_n); x_k \geq 0\}$

Economy: There are I consumers, J producers, with initial endowments of commodities $\omega = (\omega_1, \ldots, \omega_n)$.
Consumer i has a utility function $U^i(x_1^i, \ldots, x_n^i)$, $i = 1, \ldots, I$.
Producer j has a production transformation function $F^j(y_1^j, \ldots, y_n^j) = 0$,

A price system: $(p_1, \ldots, p_n)$.

A private ownership economy: Endowments and firms (producers) are owned by consumers.
Consumer i's endowment is $\omega^i = (\omega_1^i, \ldots, \omega_n^i)$, $\sum_{i=1}^I \omega^i = \omega$.
Consumer i's share of firm j is $\theta^{ij} \geq 0$, $\sum_{i=1}^I \theta^{ij} = 1$.

An allocation: $x^i = (x_1^i, \ldots, x_n^i)$, $i = 1, \ldots, I$, and $y^j = (y_1^j, \ldots, y_n^j)$, $j = 1, \ldots, J$.

A competitive equilibrium:
A combination of a price system $\bar{p} = (\bar{p}_1, \ldots, \bar{p}_n)$ and an allocation $(\{\bar{x}^i\}_{i=1,\ldots,I}, \{\bar{y}^j\}_{j=1,\ldots,J})$ such that
1. $\sum_i \bar{x}^i = \omega + \sum_j \bar{y}^j$ (feasibility condition).
2. $\bar{y}^j$ maximizes Π^j, $j = 1, \ldots, J$ and $\bar{x}^i$ maximizes U^i, subject to i's budget constraint $p_1 x_1^i + \ldots + p_n x_n^i = p_1 \omega_1^1 + \ldots + p_n \omega_n^i + \theta_{i1}\Pi^1 + \ldots + \theta_{iJ}\Pi^J$.

Existence Theorem:
Suppose that the utility functions are all quasi-concave and the production transformation functions satisfy some theoretic conditions, then a competitive equilibrium exists.

Welfare Theorems: A competitive equilibrium is efficient and an efficient allocation can be achieved as a competitive equilibrium through certain income transfers.

Constant returns to scale economies and non-substitution theorem:
Suppose there is only one nonproduced input, this input is indispensable to production, there is no joint production, and the production functions exhibits constant returns to scale. Then the competitive equilibrium price system is determined by the production side only.

10.5 Multi-person Decision Problem and Game Theory

In this chapter, we consider the situation when there are $n > 1$ persons with different objective (utility) functions; that is, different persons have different preferences over possible outcomes. There are two cases:
1. Game theory: The outcome depends on the behavior of all the persons involved. Each person has some control over the outcome; that is, each person controls certain strategic variables. Each one's utility depends on the decisions of all persons. We want to study how persons make decisions.

2. Public Choice: Persons have to make decision collectively, eg., by voting.
We consider only game theory here.

Game theory: the study of conflict and cooperation between persons with different objective functions.

Example (a 3-person game): The accuracy of shooting of A, B, C is 1/3, 2/3, 1, respectively. Each person wants to kill the other two to become the only survivor. They shoot in turn starting A.
Question: What is the best strategy for A?

10.6 Ingredients and classifications of games

A game is a collection of rules known to all players which determine what players may do and the outcomes and payoffs resulting from their choices.
The ingredients of a game:

1. Players: Persons having some influences upon possible income (decision makers).

2. Moves: decision points in the game at which players must make choices between alternatives (personal moves) and randomization points (called nature's moves).

3. A play: A complete record of the choices made at moves by the players and realizations of randomization.

4. Outcomes and payoffs: a play results in an outcome, which in turn determines the rewords to players.

Classifications of games:

1. according to number of players:
   2-person games – conflict and cooperation possibilities.
   n-person games – coalition formation possibilities in addition.
   infinite-players' games – corresponding to perfect competition in economics.

2. according to number of strategies:
   finite – strategy (matrix) games, each person has a finite number of strategies,

payoff functions can be represented by matrices.
infinite – strategy (continuous or discontinuous payoff functions) games like duopoly games.

3. according to sum of payoffs:
   0-sum games – conflict is unavoidable.
   non-zero sum games – possibilities for cooperation.

4. according to preplay negotiation possibility:
   non-cooperative games – each person makes unilateral decisions.
   cooperative games – players form coalitions and decide the redistribution of aggregate payoffs.

10.7 The extensive form and normal form of a game

Extensive form: The rules of a game can be represented by a game tree.
The ingredients of a game tree are:
1. Players
2. Nodes: they are players' decision points (personal moves) and randomization points (nature's moves).
3. Information sets of player i: each player's decision points are partitioned into information sets. An information set consists of decision points that player i can not distinguish when making decisions.
4. Arcs (choices): Every point in an information set should have the same number of choices.
5. Randomization probabilities (of arcs following each randomization points).
6. Outcomes (end points)
7. Payoffs: The gains to players assigned to each outcome.
A pure strategy of player i: An instruction that assigns a choice for each information set of player i.
Total number of pure strategies of player i: the product of the numbers of choices of all information sets of player i.

Once we identify the pure strategy set of each player, we can represent the game in normal form (also called strategic form).

1. Strategy sets for each player: $S_1 = \{s_1, \ldots, s_m\}$, $S_2 = \{\sigma_1, \ldots, \sigma_n\}$.

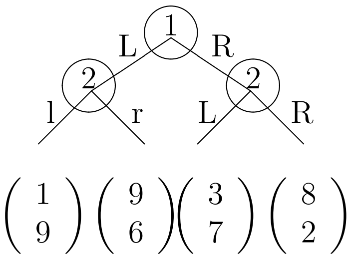
2. Payoff matrices: $\pi_1(s_i, \sigma_j) = a_{ij}$, $\pi_2(s_i, \sigma_j) = b_{ij}$. $A = [a_{ij}]$, $B = [b_{ij}]$.

Normal form:

I \ II	σ_1	$\ldots$	σ_n
s_1	(a_{11}, b_{11})	$\ldots$	(a_{1n}, b_{1n})
$\vdots$	$\vdots$	$\ddots$	$\vdots$
s_m	(a_{m1}, b_{m1})	$\ldots$	(a_{mn}, b_{mn})

10.8 Examples

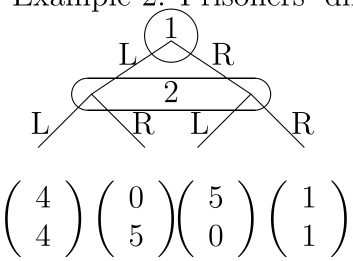
Example 1: A perfect information game



$S_1 = \{$ L, R $\}$, $S_2 = \{$ Ll, Lr, Rl, Rr $\}$.

I \ II	Ll	Lr	Rl	Rr
L	(1,9)	(9,6)	(1,9)	(9,6)
R	(3,7)*	(8,2)	(3,7)	(8,2)

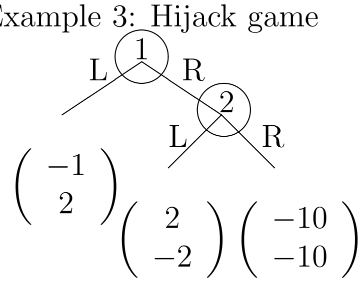
Example 2: Prisoners' dilemma game



$S_1 = \{$ L, R $\}$, $S_2 = \{$ L, R $\}$.

I \ II	L	R
L	(4,4)	(0,5)
R	(5,0)	(1,1)*

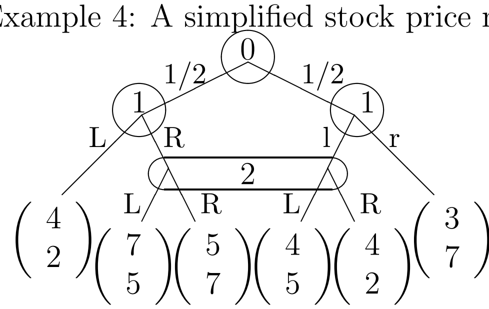
Example 3: Hijack game



$S_1 = \{$ L, R $\}$, $S_2 = \{$ L, R $\}$.

I \ II	L	R
L	(-1,2)	(-1,2)*
R	(2,-2)*	(-10,-10)

Example 4: A simplified stock price manipulation game



$S_1 = \{$ Ll, Lr, Rl, Rr $\}$, $S_2 = \{$ L, R $\}$.

I \ II	L	R
Ll	(4, 3.5)	(4, 2)
Lr	(3.5, 4.5)	(3.5, 4.5)
Rl	(5.5, 5)*	(4.5, 4.5)
Rr	(5,6)	(4,7)

Remark: Each extensive form game corresponds a normal form game. However, different extensive form games may have the same normal form.

10.9 Strategy pair and pure strategy Nash equilibrium

1. A Strategy Pair: (s_i, σ_j). Given a strategy pair, there corresponds a payoff pair (a_{ij}, b_{ij}).

2. A Nash equilibrium: A strategy pair (s_{i*}, σ_{j*}) such that $a_{i*j*} \geq a_{ij*}$ and $b_{i*j*} \geq b_{i*j}$ for all (i, j). Therefore, there is no incentives for each player to deviate from the equilibrium strategy. a_{i*j*} and b_{i*j*} are called the equilibrium payoff.

The equilibrium payoffs of the examples are marked each with a star in the normal form.

Remark 1: It is possible that a game does no have a pure strategy Nash equilibrium. Also, a game can have more than one Nash equilibria.
Remark 2: Notice that the concept of a Nash equilibrium is defined for a normal form game. For a game in extensive form (a game tree), we have to find the normal form before we can find the Nash equilibria.

10.10 Subgames and subgame perfect Nash equilibria

1. Subgame: A subgame in a game tree is a part of the tree consisting of all the nodes and arcs following a node that form a game by itself.

2. Within an extensive form game, we can identify some subgames.

3. Also, each pure strategy of a player induces a pure strategy for every subgame.

4. Subgame perfect Nash equilibrium: A Nash equilibrium is called **subgame perfect** if it induces a Nash equilibrium strategy pair for every subgame.

5. Backward induction: To find a subgame perfect equilibrium, usually we work backward. We find Nash equilibria for lowest level (smallest) subgames and replace the subgames by its Nash equilibrium payoffs. In this way, the size of the game is reduced step by step until we end up with the equilibrium payoffs.

All the equilibria, except the equilibrium strategy pair (L,R) in the hijack game, are subgame perfect.
Remark: The concept of a subgame perfect Nash equilibrium is defined only for an extensive form game.

10.10.1 Perfect information game and Zemelo's Theorem

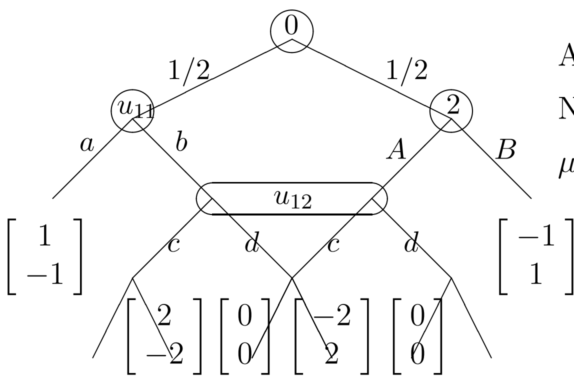
An extensive form game is called perfect information if every information set consists only one node. Every perfect information game has a pure strategy subgame perfect Nash Equilibrium.

10.10.2 Perfect recall game and Kuhn's Theorem

A local strategy at an information set $u \in U_i$: A probability distribution over the choice set at U_{ij}.
A behavior strategy: A function which assigns a local strategy for each $u \in U_i$.
The set of behavior strategies is a subset of the set of mixed strategies.

Kuhn's Theorem: In every extensive game with perfect recall, a strategically equivalent behavior strategy can be found for every mixed strategy.

However, in a non-perfect recall game, a mixed strategy may do better than behavior strategies because in a behavior strategy the local strategies are independent whereas they can be correlated in a mixed strategy.



A 2-person 0-sum non-perfect recall game.
NE is $(\mu_1^*, \mu_2^*) = (\frac{1}{2}ac \oplus \frac{1}{2}bd, \frac{1}{2}A \oplus \frac{1}{2}B)$.
μ_1^* is not a behavioral strategy.

10.10.3 Reduction of a game

Redundant strategy: A pure strategy is redundant if it is strategically identical to another strategy.
Reduced normal form: The normal form without redundant strategies.
Equivalent normal form: Two normal forms are equivalent if they have the same reduced normal form.
Equivalent extensive form: Two extensive forms are equivalent if their normal forms are equivalent.

10.11 Continuous games and the duopoly game

In many applications, S_1 and S_2 are infinite subsets of R^m and R^n Player 1 controls m variables and player 2 controls n variables (however, each player has infinite many strtategies). The normal form of a game is represented by two functions

$$\Pi^1 = \Pi^1(x;y) \quad \text{and} \quad \Pi^2 = \Pi^2(x;y), \quad \text{where} \quad x \in S_1 \subset R^m \quad \text{and} \quad y \in S_2 \subset R^n.$$

To simplify the presentation, assume that $m = n = 1$. A strategic pair is $(x, y) \in S_1 \times S_2$. A Nash equilibrium is a pair (x^*, y^*) such that

$$\Pi^1(x^*, y^*) \geq \Pi^1(x, y^*) \quad \text{and} \quad \Pi^2(x^*, y^*) \geq \Pi^2(x^*, y) \qquad \text{for all} \quad x \in S_1 \quad y \in S_2.$$

Consider the case when Π^i are continuously differentiable and Π^1 is strictly concave in x and Π^2 strictly concave in y (so that we do not have to worry about the SOC's).

Reaction functions and Nash equilibrium:
To player 1, x is his endogenous variable and y is his exogenous variable. For each y chosen by player 2, player 1 will choose a $x \in S_1$ to maximize his objective function Π^1. This relationship defines a behavioral equation $x = R^1(y)$ which can be obtained by solving the FOC for player 1, $\Pi^1_x(x; y) = 0$. Similarly, player 2 regards y as endogenous and x exogenous and wants to maximize Π^2 for a given x chosen by player 1. Player 2's reaction function (behavioral equation) $y = R^2(x)$ is obtained by solving $\Pi^2_y(x; y) = 0$. A Nash equilibrium is an intersection of the two reaction functions. The FOC for a Nash equilibrium is given by $\Pi^1_x(x^*; y^*) = 0$ and $\Pi^2_y(x^*; y^*) = 0$.

Duopoly game:
There are two sellers (firm 1 and firm 2) of a product.
The (inverse) market demand function is $P = a - Q$.
The marginal production costs are c_1 and c_2, respectively.
Assume that each firm regards the other firm's output as given (not affected by his output quantity).
The situation defines a 2-person game as follows: Each firm i controls his own output quantity q_i. (q_1, q_2) together determine the market price $P = a - (q_1 + q_2)$ which in turn determines the profit of each firm:

$$\Pi^1(q_1, q_2) = (P - c_1)q_1 = (a - c_1 - q_1 - q_2)q_1 \quad \text{and} \quad \Pi^2(q_1, q_2) = (P - c_2)q_2 = (a - c_2 - q_1 - q_2)q_2$$

The FOC are $\partial\Pi^1/\partial q_1 = a - c_1 - q_2 - 2q_1 = 0$ and $\partial\Pi^2/\partial q_2 = a - c_2 - q_1 - 2q_2 = 0$.
The reaction functions are $q_1 = 0.5(a - c_1 - q_2)$ and $q_2 = 0.5(a - c_2 - q_1)$.
The Cournot Nash equilibrium is $(q_1^*, q_2^*) = ((a - 2c_1 + c_2)/3, (a - 2c_2 + c_1)/3)$ with $P^* = (a + c_1 + c_2)/3$. (We have to assume that $a - 2c_1 + c_2,\ a - 2c_2 + c_1 \geq 0$.)

10.11.1 A simple bargaining model

Two players, John and Paul, have \$ 1 to divide between them. They agree to spend at most two days negotiating over the division. The first day, John will make an offer, Paul either accepts or comes back with a counteroffer the second day. If they cannot reach an agreement in two days, both players get zero. John (Paul) discounts payoffs in the future at a rate of α (β) per day.

A subgame perfect equilibrium of this bargaining game can be derived using backward induction.
1. On the second day, John would accept any non-negative counteroffer made by Paul. Therefore, Paul would make proposal of getting the whole \$ 1 himself and John would get \$ 0.
2. On the first day, John should make an offer such that Paul gets an amount equivalent to getting \$ 1 the second day, otherwise Paul will reject the offer. Therefore, John will propose of $1 - \beta$ for himself and β for Paul and Paul will accept the offer.

An example of a subgame non-perfect Nash equilibrium is that John proposes of getting 1-0.5β for himself and 0.5β for Paul and refuses to accept any counteroffer made by Paul. In this equilibrium, Paul is threatened by John's incredible threat and accepts only one half of what he should have had in a perfect equilibrium.

10.12 2-person 0-sum game

1. $B = -A$ so that $a_{ij} + b_{ij} = 0$.

2. Maxmin strategy: If player 1 plays s_i, then the minimum he will have is $\min_j a_{ij}$, called the security level of strategy s_i. A possible guideline for player 1 is to choose a strategy such that the security level is maximized: Player 1 chooses s_{i*} so that $\min_j a_{i*j} \geq \min_j a_{ij}$ for all i. Similarly, since $b_{ij} = -a_{ij}$, Player 2 chooses σ_j* so that $\max_i a_{ij*} \leq \max_i a_{ij}$ for all j.

3. Saddle point: If $a_{i*j*} = \max_i \min_j a_{ij} = \min_j \max_i a_{ij}$, then (s_{i*}, σ_{j*}) is called a saddle point. If a saddle point exists, then it is a Nash equilibrium.

$$A_1 = \begin{pmatrix} 2 & 1 & 4 \\ -1 & 0 & 6 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   In example A_1, $\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = 1$ (s_1, σ_2) is a saddle point and hence a Nash equilibrium. In A_2, $\max_i \min_j a_{ij} = 0 \neq \min_j \max_i a_{ij} = 1$ and no saddle point exists. If there is no saddle points, then there is no pure strategy equilibrium.

4. Mixed strategy for player i: A probability distribution over S_i. $p = (p_1, \ldots, p_m)$, $q = (q_1, \ldots, q_n)'$. (p, q) is a mixed strategy pair. Given (p, q), the expected payoff of player 1 is pAq. A mixed strategy Nash equilibrium (p^*, q^*) is such that $p^*Aq^* \geq pAq^*$ and $p^*Aq^* \leq p^*Aq$ for all p and all q.

5. Security level of a mixed strategy: Given player 1's strategy p, there is a pure strategy of player 2 so that the expected payoff to player 1 is minimized, just as in the case of a pure strategy of player 1.

$$t(p) \equiv \min_j \{\sum_i p_i a_{i1}, \ldots, \sum_i p_i a_{in}\}.$$

   The problem of finding the maxmin mixed strategy (to find p^* to maximize $t(p)$) can be stated as

$$\max_p t \quad \text{subj. to} \quad \sum_i p_i a_{i1} \geq t, \ldots, \sum_i p_i a_{in} \geq t, \;\; \sum_i p_i = 1.$$

6. Linear programming problem: The above problem can be transformed into a linear programming problem as follows: (a) Add a positive constant to each element of A to insure that $t(p) > 0$ for all p. (b) Define $y_i \equiv p_i/t(p)$ and

replace the problem of $\max t(p)$ with the problem of $\min 1/t(p) = \sum_i y_i$. The constraints become $\sum_i y_i a_{i1} \geq 1, \ldots, \sum_i y_i a_{in} \geq 1$.

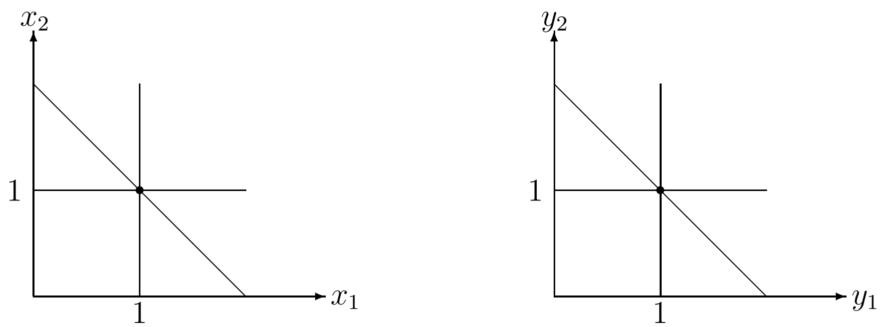
$$\min_{y_1,\ldots,y_m \geq 0} y_1 + \ldots + y_m \quad \text{subj. to} \sum_i y_i a_{i1} \geq 1, \ldots, \sum_i y_i a_{in} \geq 1$$

7. Duality: It turns out that player 2's minmax problem can be transformed similarly and becomes the dual of player 1's linear programming problem. The existence of a mixed strategy Nash equilibrium is then proved by using the duality theorem in linear programming.

Example (tossing coin game): $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
To find player 2's equilibrium mixed strategy, we solve the linear programming problem:

$$\max_{x_1,x_2 \geq 0} x_1 + x_2 \quad \text{subj. to} \quad x_1 \leq 1 \;\; x_2 \leq 1.$$

The solution is $x_1 = x_2 = 1$ and therefore the equilibrium strategy for player 2 is $q_1^* = q_2^* = 0.5$.



Player 1's equilibrium mixed strategy is obtained by solving the dual to the linear programming problem:

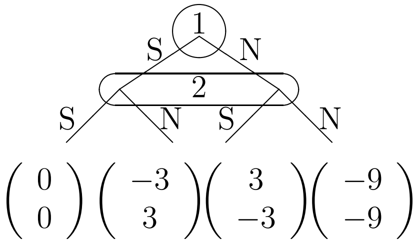
$$\min_{y_1,y_2 \geq 0} y_1 + y_2 \quad \text{subj. to} \quad y_1 \geq 1 \;\; y_2 \geq 1.$$

The solution is $p_1^* = p_2^* = 0.5$.

10.13 Mixed strategy equilibria for non-zero sum games

The idea of a mixed strategy equilibrium is also applicable to a non-zero sum game. Similar to the simplex algorism for the 0-sum games, there is a Lemke algorism.

Example (Game of Chicken)

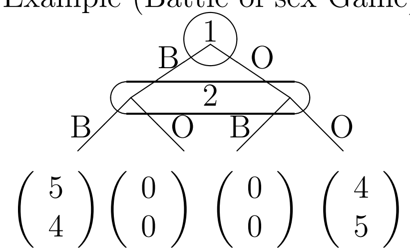


I \ II	Swerve	Don't
Swerve	(0,0)	(-3,3)*
Don't	(3,-3)*	(-9,-9)

$S_1 = \{$ S, N $\}$, $S_2 = \{$ S, N $\}$.

There are two pure strategy NE: (S, N) and (N, S).
There is also a mixed strategy NE. Suppose player 2 plays a mixed strategy $(q, 1-q)$. If player 1 plays S, his expected payoff is $\Pi^1(S) = 0q + (-3)(1-q)$. If he plays N, his expected payoff is $\Pi^1(N) = 3q + (-9)(1-q)$. For a mixed strategy NE, $\Pi^1(S) = \Pi^1(N)$, therefore, $q = \frac{2}{3}$.
The mixed strategy is symmetrical: $(p_1^*, p_2^*) = (q_1^*, q_2^*) = (\frac{2}{3}, \frac{1}{3})$.

Example (Battle of sex Game)



I \ II	Ball game	Opera
Ball game	(5,4)*	(0,0)
Opera	(0,0)	(4,5)*

$S_1 = \{$ B, O $\}$, $S_2 = \{$ B, O $\}$.

There are two pure strategy NE: (B, B) and (O, O).
There is also a mixed strategy NE. Suppose player 2 plays a mixed strategy $(q, 1-q)$. If player 1 plays B, his expected payoff is $\Pi^1(B) = 5q + (0)(1-q)$. If he plays O, his expected payoff is $\Pi^1(O) = 0q + (4)(1-q)$. For a mixed strategy NE, $\Pi^1(B) = \Pi^1(O)$, therefore, $q = \frac{4}{9}$.
The mixed strategy is: $(p_1^*, p_2^*) = (\frac{5}{9}, \frac{4}{9})$ and $(q_1^*, q_2^*) = (\frac{4}{9}, \frac{5}{9})$.

10.14 Cooperative Game and Characteristic form

2-person 0-sum games are strictly competitive. If player 1 gains \$ 1, player 2 will loss \$ 1 and therefore no cooperation is possible. For other games, usually some cooperation is possible. The concept of a Nash equilibrium is defined for the situation when no explicit cooperation is allowed. In general, a Nash equilibrium is not efficient (not Pareto optimal). When binding agreements on strategies chosen can be contracted before the play of the game and transfers of payoffs among players after a play of the game is possible, players will negotiate to coordinate their strategies and redistribute the payoffs to achieve better results. In such a situation, the determination of strategies is not the key issue. The problem becomes the formation of coalitions and the distribution of payoffs.

Characteristic form of a game:
The player set: $N = \{1, 2, \ldots, n\}$.
A coalition is a subset of N: $S \subset N$.
A characteristic function v specifies the maximum total payoff of each coalition.
Consider the case of a 3-person game. There are 8 subsets of $N = \{1, 2, 3\}$, namely, $\phi, (1), (2), (3), (12), (13), (23), (123)$.
Therefore, a characteristic form game is determined by 8 values $v(\phi), v(1), v(2), v(3), v(12), v(13), v(2$
Super-additivity: If $A \cap B = \phi$, then $v(A \cup B) \geq v(A) + v(B)$.
An imputation is a payoff distribution (x_1, x_2, x_3).
Individual rationality: $x_i \geq v(i)$.
Group rationality: $\sum_{i \in S} x_i \geq v(S)$.
Core C: the set of imputations that satisfy individual rationality and group rationality for all S.

Marginal contribution of player i in a coalition $S \cup i$: $v(S \cup i) - v(S)$
Shapley value of player i is an weighted average of all marginal contributions

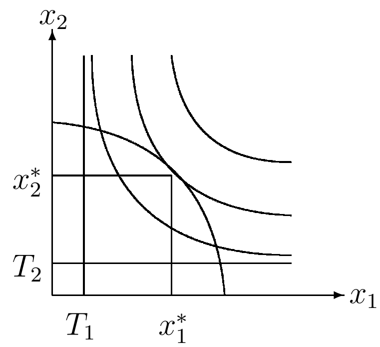
$$\pi_i = \sum_{S \subset N} \frac{|S|!(n - |S| - 1)!}{n!} [v(S \cup i) - v(S)].$$

Example: $v(\phi) = v(1) = v(2) = v(3) = 0, v(12) = v(13) = v(23) = 0.5, v(123) = 1$.
$C = \{(x_1, x_2, x_3),\ x_i \geq 0, x_i + x_j \geq 0.5, x_1 + x_2 + x_3 = 1\}$. Both (0.3, 0.3, 0.4) and (0.2, 0.4, 0.4) are in C.
The Shapley values are $(\pi_1, \pi_2, \pi_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Remark 1: The core of a game can be empty. However, the Shapley values are uniquely determined.
Remark 2: Another related concept is the von-Neumann Morgenstern solution. See CH 6 of Intriligator's *Mathematical Optimization and Economic Theory* for the motivations of these concepts.

10.15 The Nash bargaining solution for a nontransferable 2-person cooperative game

In a nontransferable cooperative game, after-play redistributions of payoffs are impossible and therefore the concepts of core and Shapley values are not suitable. For the case of 2-person games, the concept of Nash bargaining solutions are useful.
Let $F \subset R^2$ be the feasible set of payoffs if the two players can reach an agreement and T_i the payoff of player i if the negotiation breaks down. T_i is called the threat point of player i. The Nash bargaining solution (x_1^*, x_2^*) is defined to be the solution to the following problem:



$$\max_{(x_1,x_2)\in F} (x_1 - T_1)(x_2 - T_2)$$

See CH 6 of Intriligator's book for the motivations of the solution concept.

10.16 Problems

1. Consider the following two-person 0-sum game:

I \ II	σ_1	σ_2	σ_3
s_1	4	3	-2
s_2	3	4	10
s_3	7	6	8

   (a) Find the max min strategy of player I $s_{\max\min}$ and the min max strategy of player II $\sigma_{\min\max}$.

   (b) Is the strategy pair $(s_{\max\min}, \sigma_{\min\max})$ a Nash equilibrium of the game?

   (c) What are the equilibrium payoffs?

2. Find the maxmin strategy ($s_{\max\min}$) and the minmax strategy ($\sigma_{\min\max}$) of the following two-person 0-sum game:

I \ II	σ_1	σ_2
s_1	-3	6
s_2	8	-2
s_3	6	3

   Is the strategy pair ($s_{\max\min}$, $\sigma_{\min\max}$) a Nash equilibrium? If not, use simplex method to find the mixed strategy Nash equilibrium.

3. Find the (mixed strategy) Nash Equilibrium of the following two-person game:

I \ II	H	T
H	(-2, 2)	(2, -1)
T	(2, -2)	(-1,2)

4. Suppose that two firms producing a homogenous product face a linear demand curve $P = a - bQ = a - b(q_1 + q_2)$ and that both have the same constant marginal costs c. For a given quantity pair (q_1, q_2), the profits are $\Pi_i = q_i(P - c) = q_i(a - bq_1 - bq_2 - c)$, $i = 1, 2$. Find the Cournot Nash equilibrium output of each firm.

5. Suppose that in a two-person cooperative game without side payments, if the two players reach an agreement, they can get (Π_1, Π_2) such that $\Pi_1^2 + \Pi_2 = 47$

and if no agreement is reached, player 1 will get $T_1 = 3$ and player 2 will get $T_2 = 2$.

(a) Find the Nash solution of the game.

(b) Do the same for the case when side payments are possible. Also answer how the side payments should be done?

6. A singer (player 1), a pianist (player 2), and a drummer (player 3) are offered \$ 1,000 to play together by a night club owner. The owner would alternatively pay \$ 800 the singer-piano duo, \$ 650 the piano drums duo, and \$ 300 the piano alone. The night club is not interested in any other combination. Howeover, the singer-drums duo makes \$ 500 and the singer alone gets \$ 200 a night in a restaurant. The drums alone can make no profit.

   (a) Write down the characteristic form of the cooperative game with side payments.

   (b) Find the Shapley values of the game.

   (c) Characterize the core.