APEC Math Review

Part 7 Linear Algebra

Ling Yao

August, 2020

University of Minnesota

1. Solving linear systems

Linear systems

A linear system of m equations in n unknowns takes the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m2}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Gaussian elimination

Linear systems can be solved by substitutions and eliminations.
e.g.

$$\begin{aligned} x_1 - 0.4x_2 - 0.3x_3 &= 130 \\ -0.2x_1 + 0.88x_2 - 0.14x_3 &= 74 \\ -0.5x_1 - 0.2x_2 + 0.95x_3 &= 95 \end{aligned}$$

$\rightarrow$

$$\begin{aligned} x_1 - 0.4x_2 - 0.3x_3 &= 130 \\ 0.8x_2 - 0.2x_3 &= 100 \\ 0.7x_3 &= 210 \end{aligned}$$

$\rightarrow$

$$\begin{aligned} x_1 &= 300 \\ x_2 &= 200 \\ x_3 &= 300 \end{aligned}$$

Elementary row operations

Alternatively, we can use elementary row operations to solve the linear system.
First, obtain the **augmented matrix**

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{12} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m2} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Next, do the following **elementary row operations** until the matrix is in the **row echelon form.**

- interchange two rows
- change a row by adding to it a multiple of another row
- multiply each element in a row by the same nonzero number

Elementary row operations

e.g.

$$\left(\begin{array}{ccc|c} 1 & -0.4 & -0.3 & 130 \\ -0.2 & 0.88 & -0.14 & 74 \\ -0.5 & -0.2 & 0.95 & 95 \end{array}\right)$$

$\rightarrow$

$$\left(\begin{array}{ccc|c} 1 & -0.4 & -0.3 & 130 \\ 0 & 0.8 & -0.2 & 100 \\ 0 & 0 & 0.7 & 210 \end{array}\right)$$

$\rightarrow$

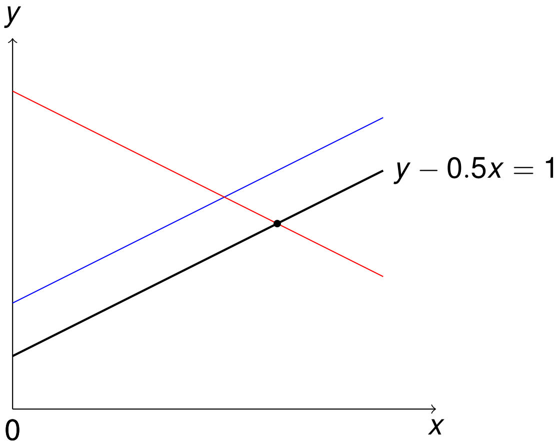
$$\left(\begin{array}{ccc|c} 1 & -0.4 & -0.3 & 130 \\ 0 & 1 & -0.25 & 125 \\ 0 & 0 & 1 & 300 \end{array}\right)$$

$\rightarrow$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 300 \\ 0 & 1 & 0 & 200 \\ 0 & 0 & 1 & 300 \end{array}\right)$$

The existence of solution(s)

Consider a two equations system.



How do we know whether a linear system has one solution, no solution or infinitely many solutions?

Rank

The **rank** of a matrix is the number of nonzero rows in its row echelon form.

- $rank\mathbf{A} = rank\mathbf{A}' \leq min(\#rows, \#columns)$
- $rank\mathbf{AB} \leq min(rank\mathbf{A}, rank\mathbf{B})$
- $rank\mathbf{A} = rank\mathbf{A}'\mathbf{A} = rank\mathbf{AA}'$
- A matrix is **full rank** if the rank equals to the number of columns.

Let **A** be the coefficient matrix and $\hat{\mathbf{A}}$ be the corresponding augmented matrix,

- $rank\mathbf{A} \leq$ number of rows of **A**
- $rank\mathbf{A} \leq$ number of columns of **A**
- $rank\hat{\mathbf{A}} \geq rank\mathbf{A}$

The existence of solution(s)

A system of linear equations with coefficient matrix **A** and augmented matrix $\hat{\mathbf{A}}$ has a solution iif

$$rank\hat{\mathbf{A}} = rank\mathbf{A}$$

That is, no augmented matrix like this form

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & b_m \end{array}\right)$$

Exercise

(Simon & Blume Exercise 7.18)
For what values of the parameter a does the following system of equations have a solution?

$$\begin{aligned} 6x + y &= 7 \\ 3x + y &= 4 \\ -6x - 2y &= a \end{aligned}$$

The existence of solution(s)

A linear system has infinitely many solutions if

$$\textit{number of rows of } \mathbf{A} < \textit{number of columns of } \mathbf{A}.$$

A linear system has one and only one solution for every choice of right-hand side $b_1, \cdots, b_m$ iif

$$\textit{number of rows of } \mathbf{A} = \textit{number of columns of } \mathbf{A} = \textit{rank}\,\mathbf{A}$$

Such a coefficient matrix **A** is called a **nonsingular square matrix.**

Linear independence

A **homogeneous** system, which has $b_i = 0$ for all i,

$$\mathbf{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0}$$

is guaranteed to have at least one solution: $x_i = 0$ for all i.

But iif there is a nonzero solution (which means there is infinitely more), each column vectors $\mathbf{a_1}, \cdots, \mathbf{a_n}$ in A are **linearly dependent** . This means at least one of the column vectors can be written as **linear combination** of the others

A set of vector is **linearly independent** if and only if the only solution is the zero solution.

Linear independence

If **A** is square, having a nonzero solution means **A** is singular.
If **A** has more columns than rows (i.e. **A** is **short rank**), then the columns must linearly dependent.

Linear implicit function

Sometimes the equations are implicit form. Let $x_1, \cdots, x_k$ and $x_{k+1}, \cdots, x_n$ be a partition of the n variables in the linear system into endogenous and exogenous variables, respectively. There is, for each choice of values for the endogenous variables, a unique set of values for the exogenous variables, which solves the system iif

1. $k = m$ (# endogenous variables = # equations) and
2. the rank of the coefficient matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{12} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m2} & a_{m2} & \cdots & a_{mk} \end{pmatrix}$$

corresponding to the endogenous variables, is k.

Exercise: structural and reduced form

(Simon & Blume, Exercise 7.25)
For the following system, separate the variables into exogenous and endogenous ones so that each choice of values of the exogenous variables determines unique values for the endogenous variables. Also, find an explicit formula for the endogenous variables in terms of the exogenous ones.

$$\begin{aligned} x + 2y + z - w &= 1 \\ 3x - y - z - 3w &= 3 \\ y + z + w &= 0 \end{aligned}$$

2. Square matrices

Square matrices

Linear systems can also be solved with matrix inversion if the number of unknowns is the same as the number of equations, that is, the coefficient matrix is square and nonsingular.

Inversion

Let $\mathbf{B} = \mathbf{A}^{-1}$ be the inverse of a full-rank $k \times k$ matrix **A**. The matrices satisfy

$$\mathbf{A}\mathbf{B} = \mathbf{I}_k$$

If an $n \times n$ matrix **A** is invertible, then it is nonsingular, and the unique solution to the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Inversion

The following statements about a $n \times n$ square matrix $\mathbf{A}$ are equivalent

- $\mathbf{A}$ is invertible
- $\mathbf{A}$ is nonsingular
- A has maximal rank n (full rank)
- every system $\mathbf{Ax} = \mathbf{b}$ has one and only one solution for every $\mathbf{b}$

Determinant

A $n \times n$ square matrix $\mathbf{A}$ is nonsingular if and only if its **determinant** is nonzero.

$$det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \ or \ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$- a_{13}a_{22}a_{31} - a_{21}a_{12}a_{33} - a_{11}a_{23}a_{32}$$

- $det\,\mathbf{A}^T = det\,\mathbf{A}$
- $det\,(\mathbf{A} \cdot \mathbf{B}) = det\,\mathbf{A} \cdot det\,\mathbf{B}$
- $det\,(\mathbf{A} + \mathbf{B}) \neq det\,\mathbf{A} + det\,\mathbf{B}$

What will happen if $\mathbf{A}$ is not full rank?

Inversion

Let $\mathbf{A}$ and $\mathbf{B}$ be square invertible matrices. Then,

1. $det\, \mathbf{A}$ and $det\, \mathbf{B}$ are nonzero
2. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
3. $(\mathbf{A}^{\mathbf{T}})^{-1} = (\mathbf{A}^{-1})^{\mathbf{T}}$
4. $\mathbf{AB}$ is invertible, and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Inversion

Let **A** be a nonsingular matrix,

$$\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}} \cdot adj\,\mathbf{A},$$

where *adj* **A** is a $n \times n$ square matrix in which the element on the *i*th row and *j* the column is

$$(-1)^{i+j} \times det(submatrix\ of\ \mathbf{A}\ without\ the\ ith\ row\ and\ the\ jth\ column)$$

Exercise: matrix inversion

Invert the following matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 1 & 0 & 8 \end{pmatrix}$$

Cramer's rule

The unique solution $\mathbf{x} = (x_1, \cdots, x_n)$ of the $n \times n$ system $\mathbf{Ax} = \mathbf{b}$ is

$$x_i = \frac{det\, \mathbf{B}_i}{det\, \mathbf{A}} \quad \textit{for } i = 1, \cdots, n,$$

where $\mathbf{B}_i$ is the matrix $\mathbf{A}$ with the right-hand side $\mathbf{b}$ replacing the ith column of $\mathbf{A}$.

Exercise: Cramer's rule

Use the Cramer's rule to calculate x_3 in the following system:

$$\begin{pmatrix} 1 & 1 & 1 \\ 12 & 2 & -3 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ -4 \end{pmatrix}$$

Trace

The trace of a $k \times k$ square matrix $\mathbf{A}$ is the sum of its diagonal elements

$$tr(\mathbf{A}) = \sum_{i=1}^{k} a_{ii}$$

Properties

- $tr(c\mathbf{A}) = c\, tr(\mathbf{A})$
- $tr(\mathbf{A}') = tr(\mathbf{A})$
- $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$
- $tr(\mathbf{I}_k) = k$
- for $k \times r$ $\mathbf{A}$ and $r \times k$ $\mathbf{B}$, $tr(\mathbf{AB}) = tr(\mathbf{BA})$

Characteristic equation

For a square matrix **A**, we can define a set of equations

$$\mathbf{A}\mathbf{c} = \lambda \mathbf{c}$$

The pairs of solutions $(\mathbf{c}, \lambda)$ are the **characteristic vectors (eigenvectors) c** and **characteristic roots (eigenvalues)** λ**.**
The equations imply that

$$\mathbf{A}\mathbf{c} = \lambda \mathbf{I}\mathbf{c} \Leftrightarrow (\mathbf{A} - \lambda \mathbf{I})\mathbf{c} = \mathbf{0}$$

This homogeneous system has a nonzero solution only if $(\mathbf{A} - \lambda \mathbf{I})$ is singular and has zero determinant.

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

This polynomial of λ is the characteristic equation.

Exercise: characteristic equation

Solve for the characteristic roots for the following matrix.

$$\mathbf{A} = \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix}$$

Characteristic vectors

With the characteristic roots, we can solve for the characteristic vectors corresponding to each characteristic roots using

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{c} = \mathbf{0}$$

$(\mathbf{A} - \lambda \mathbf{I})$ is singular so there exists a non-zero solution. In fact, there are infinite nonzero solutions, just pick one that is nonzero. Normalize it so that $\mathbf{c}'\mathbf{c} = \mathbf{1}$.

Solve for the two characteristic vectors in the last example.

Some useful results

- A $k \times k$ matrix has k distinct characteristic vectors that are orthogonal to each other.
- The rank of a symmetric matrix is the number of nonzero characteristic roots it contains.
- The trace of a matrix equals the sum of its characteristic roots.
- The determinant of a matrix equals the product of its characteristic roots.
- If $\mathbf{A}^{-1}$ exists,then the characteristic roots of $\mathbf{A}^{-1}$ are the reciprocals of those of $\mathbf{A}$ and the characteristic vectors are the same.

Decomposition

Define

$$\mathbf{C} = \begin{pmatrix} \mathbf{c_1} & \mathbf{c_2} & \cdots & \mathbf{c_k} \end{pmatrix}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix}$$

For each k

$$\mathbf{A}\mathbf{c_k} = \lambda_\mathbf{k}\mathbf{c_k}$$

So

$$\mathbf{AC} = \mathbf{C}\boldsymbol{\Lambda}$$

Decomposition

Because the vectors are orthogonal and $\mathbf{c_i'c_i} = \mathbf{1}$

$$\mathbf{C'C} = \mathbf{I}$$

So

$$\mathbf{C'} = \mathbf{C^{-1}},$$

$$\mathbf{CC'} = \mathbf{CC^{-1}} = \mathbf{I}.$$

The diagonalization of a matrix $\mathbf{A}$ is

$$\mathbf{C'AC} = \mathbf{C'C\Lambda} = \mathbf{I\Lambda} = \mathbf{\Lambda}$$

The spectral decomposition of $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{C\Lambda C'} = \sum_{i=1}^{k} \lambda_i \mathbf{c_i c_i'}$$

Definite matrices

How do we know if a symmetric matrix $\mathbf{A}$ is positive or negative (semi)definite or indefinite?

We know a symmetric matrix can be decomposed into

$$\mathbf{A} = \mathbf{C\Lambda C}'$$

Therefore,

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{C\Lambda C}'\mathbf{x} \quad \textit{for any } \mathbf{x}$$

Let $\mathbf{y} = \mathbf{C}'\mathbf{x}$. Then

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = \sum_{i=1}^{N} \lambda_i y_i^2$$

Definite matrices

- **A** is positive (negative) definite if and only if all the characteristic roots of **A** are positive (negative).
- **A** is positive (negative) semidefinite if and only if all the characteristic roots are ≥ 0 (≤ 0).
- **A** is indefinite if and only if **A** has both positive and negative characteristic roots.

Some results related to definiteness

If $\mathbf{X}$ is any $N \times K$ matrix, the symmetric $K \times K$ matrix $\mathbf{X}'\mathbf{X}$ is positive definite, if and only if $N \geq K$ and $rank(\mathbf{X}) = K$.

To see this, define a $N \times 1$ vector

$$\mathbf{p} = \mathbf{Xz}$$

for any $K \times 1$ vector $\mathbf{z} \neq \mathbf{0}$. Then

$$\mathbf{z}'\mathbf{X}'\mathbf{X}\mathbf{z} = \mathbf{p}'\mathbf{p} = \sum_{i=1}^{N} p_i^2 \geq 0$$

$$rank(\mathbf{X}'\mathbf{X} = K) \Leftrightarrow rank(\mathbf{X} = K)$$

$\Leftrightarrow$ There can exist no $\mathbf{z} \neq \mathbf{0}$ such that $\mathbf{p} = \mathbf{Xz} = \mathbf{0}$ (Recall the rank condition for a homogeneous system to have nonzero solution).
$\Leftrightarrow \sum_{i=1}^{N} p_i^2 \neq 0$
$\Leftrightarrow \mathbf{z}'\mathbf{X}'\mathbf{X}\mathbf{z} = \mathbf{p}'\mathbf{p} > 0$ (matrix $\mathbf{X}'\mathbf{X}$ is positive definite)

Exercise

Using a similar proof as in the last slide, show that if $\mathbf{X}$ is any $N \times K$ matrix, the symmetric $N \times N$ matrix $\mathbf{X}\mathbf{X}'$ is positive semidefinite if and only if $N \geq K$ and $rank(\mathbf{X}) = K$.

Another way to test definiteness: leading principle minors

Let **A** be an $N \times N$ matrix. The Kth order principal submatrix of **A** obtained by deleting the last $N - K$ rows and the last $N - K$ columns is called the Kth order **leading principal submatrix.** Its determinant is call the Kth order **leading principal minor**.

E.g. for a 3×3 matrix, the three leading principal minors are

$$\begin{vmatrix} a_{11} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Another way to test definiteness: leading principle minors

Let **A** be an $N \times N$ matrix.

- **A** is positive definite (semidefinite) if and only if all its N leading principal minors > 0 (≥ 0).
- **A** is negative definite if and only if its N leading principle minors alternate in sign as follows:

$$|\mathbf{A_1}| < 0, \quad |\mathbf{A_2}| > 0, \quad |\mathbf{A_3}| < 0, \quad \textit{etc.}$$

  (Negative semidefinte if the inequalities are weak.)
- **A** is indefinite if the leading principal minors follow other patterns.

3. Matrix calculus

System of equations

We have reviewed functions with one dependent variable. Now we turn to functions with several dependent variables. A function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is

$$\begin{aligned} q_1 &= f_1(x_1, \cdots, x_n) \\ q_2 &= f_2(x_1, \cdots, x_n) \\ &\vdots \\ q_m &= f_m(x_1, \cdots, x_n) \end{aligned}.$$

We can write this as

$$F(\mathbf{x}) = \left(f_1(x_1, \cdots, x_n),\ f_2(x_1, \cdots, x_n), ..., f_m(x_1, \cdots, x_n)\right).$$

First derivative

The Jacobian matrix is $m \times n$

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Linear functions

In a set of linear functions

$$\mathbf{y} = \mathbf{A}\mathbf{x},$$

each element y_i of $\mathbf{y}$ is

$$y_i = \mathbf{a_i'}\mathbf{x},$$

where $\mathbf{a_i'}$ is the ith row of $\mathbf{A}$. Therefore,

$$\frac{\partial y_i}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}}(\mathbf{a_i'}\mathbf{x}) = \mathbf{a_i} = \textit{transpose of ith row of}\,\mathbf{A},$$

and

$$\begin{pmatrix} (\frac{\partial y_1}{\partial \mathbf{x}})' \\ (\frac{\partial y_2}{\partial \mathbf{x}})' \\ \vdots \\ (\frac{\partial y_m}{\partial \mathbf{x}})' \end{pmatrix} = \begin{pmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \\ \vdots \\ \mathbf{a_m'} \end{pmatrix}$$

Collecting all terms, we find that $\frac{\partial}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x}) = \mathbf{A}$

Exercise

Show that $\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{a_i}) = \mathbf{a_i}$ and $\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}') = \mathbf{A}'$.

Product rule

Let **a** and **b** be $n \times 1$ vectors. Both are functions of the vector **x**. Show that

$$\frac{\partial \mathbf{a}'\mathbf{b}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}'}{\partial \mathbf{x}}\mathbf{b} + \frac{\partial \mathbf{b}'}{\partial \mathbf{x}}\mathbf{a}$$

Quadratic form

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij}$$

Notice that $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{A}'\mathbf{x}$. By the product rule,

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{I_k})\mathbf{A}\mathbf{x} + \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}')\mathbf{x} = \mathbf{I_k}\mathbf{A}\mathbf{x} + \mathbf{A}'\mathbf{x} = (\mathbf{A} + \mathbf{A}')\mathbf{x}$$

$$\frac{\partial^2}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \mathbf{A} + \mathbf{A}'$$

Approximation by differentials

$$F(\mathbf{x}^* + \Delta\mathbf{x}) - F(\mathbf{x}^*) \approx \underbrace{\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}^*) \end{bmatrix}}_{\text{Jacobian derivative}} \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix}$$

Application: the Delta method

For a linear econometric model, we can estimate a vector of coefficients **b** that has the mean of β and variance-covariance matrix ($Asy.Var[\mathbf{b}]$).

$$\mathbf{b} \overset{a}{\sim} N\Big[\beta,\, Asy.Var[\mathbf{b}]\Big]$$

We can estimate $Asy.Var[\mathbf{b}]$. But sometimes we are interested in the variance of a function of **b**, $f(\mathbf{b})$.

Application: the Delta method

For example (Greene, page 80),

$$ln(\frac{Q}{Pop})_t = \beta_1 + \beta_2 lnP_{G,t} + \beta_3 ln(\frac{Income}{Pop})_t + \beta_4 lnP_{nc,t} + \beta_5 lnP_{uc,t} + \gamma ln(\frac{G}{Pop})_{t-1} + \varepsilon_t$$

b_2 and b_3 are the estimates for the short-run price and income elasticities β_2 and β_3 , but we are also interested in the long term elasticities $\phi_2 = \beta_2/(1-\gamma)$, $\phi_3 = \beta_3/(1-\gamma)$.

Application: the Delta method

The linear approximation tells us

$$f(\mathbf{b}) \approx f(\beta) + \frac{\partial f(\beta)}{\partial \beta}(\mathbf{b} - \beta)$$

Recall that $Var(ax + b) = a^2 Var(x)$. Similarly,

$$Asy.Var[f(\mathbf{b})] = \left(\frac{\partial f(\beta)}{\partial \beta}\right) Asy.Var[\mathbf{b}] \left(\frac{\partial f(\beta)}{\partial \beta}\right)'$$

Application: the Delta method

In our example,

$$\frac{\partial \phi_2}{\partial \boldsymbol{\beta}} = \begin{pmatrix} 0 & \frac{1}{1-\gamma} & 0 & 0 & 0 & \frac{\beta_2}{(1-\gamma)^2} \end{pmatrix}$$

Write out the Jacobian for ϕ_3.

Finally we can estimate the variance for the long term elasticities ϕ_2 and ϕ_3:

$$Asy.Var[\phi_2] = \left(\frac{\partial \phi_2}{\partial \boldsymbol{\beta}}\right) Asy.Var[\mathbf{b}] \left(\frac{\phi_2}{\partial \boldsymbol{\beta}}\right)'$$

$$Asy.Var[\phi_3] = \left(\frac{\partial \phi_3}{\partial \boldsymbol{\beta}}\right) Asy.Var[\mathbf{b}] \left(\frac{\phi_3}{\partial \boldsymbol{\beta}}\right)'$$

4. Matrix representation of regressions

Application: deriving the OLS estimator

To estimate a K variables linear model with a sample of size N.

$$\mathbf{Y}_{N\times 1} = \mathbf{X}_{N\times K}\boldsymbol{\beta}_{K\times 1} + \mathbf{e}_{N\times 1}$$

Choose the $\hat{\beta}$ that minimizes the sum of squared prediction error

$$\hat{\mathbf{e}}'\hat{\mathbf{e}} = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

The normal equation

The first order condition gives us the normal equation:

$$\mathbf{X}'\mathbf{X}\hat{\beta} - \mathbf{X}'\mathbf{Y} = \mathbf{0}$$

This means

$$-\mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = -\mathbf{X}'\hat{\mathbf{e}} = \mathbf{0}$$

Q: In order to reach the solution

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y},$$

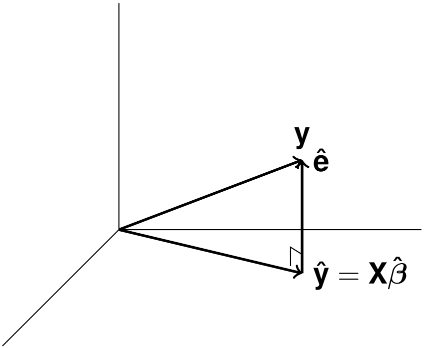
what is assumed for the **X** for the last step (rank and inner product)?

Q: Is the second order sufficient condition satisfied?

OLS from vector space view

Given a $n \times 1$ vector $\mathbf{y}$ and a $n \times k$ matrix $\mathbf{X}$, our goal is to find a $k \times 1$ vector $\hat{\beta}$ such that $\mathbf{y}$ is a linear combination of columns in $\mathbf{X}$ plus a difference:

$$\mathbf{y} = \mathbf{X}\hat{\beta} + \hat{\mathbf{e}}$$



We want the difference $\hat{\mathbf{e}}$ to be as short as possible, so $\hat{\mathbf{e}}$ should be orthorgnal to $\mathbf{X}\hat{\beta}$

$$\begin{aligned}\hat{\beta}'\mathbf{X}' \cdot \hat{\mathbf{e}} = 0 \\ &= \hat{\beta}'\mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\beta}) \\ &= \hat{\beta}[\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\beta}]\end{aligned}$$

$$\Rightarrow \mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\hat{\beta}$$

The centering matrix

In statistics, there are many cases where we need to transform the data to **deviation from their mean.** We can do it by multiplying the data matrix to a "centering matrix".
First, define a $N \times N$ matrix of $\frac{1}{N}$,

$$\frac{1}{N}\mathbf{i}\mathbf{i}' = \frac{1}{N}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \vdots & & & \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix}$$

Then define a vector of means

$$\mathbf{i}\bar{x} = \frac{1}{N}\mathbf{i}\mathbf{i}'\mathbf{x} = \begin{pmatrix} \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \vdots & & & \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{pmatrix}$$

The centering matrix

The vector of deviations can be written as

$$\mathbf{x} - \mathbf{i}\bar{x} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_N - \bar{x} \end{pmatrix} = \mathbf{x} - \frac{1}{N}\mathbf{i}\mathbf{i}'\mathbf{x} = (\mathbf{I} - \frac{1}{N}\mathbf{i}\mathbf{i}')\mathbf{x} = \mathbf{M^0}\mathbf{x}$$

The sum of deviations can be written as

$$\sum_{i=1}^{N}(x_i - x)^2 = (\mathbf{x} - \mathbf{i}\bar{x})'(\mathbf{x} - \mathbf{i}\bar{x}) = (\mathbf{M^0}\mathbf{x})'(\mathbf{M^0}\mathbf{x}) = \mathbf{x}'\mathbf{M^0}'\mathbf{M^0}\mathbf{x}$$

Show that

1. $\mathbf{M^0}\mathbf{i} = \mathbf{0}$
2. $\mathbf{M^0}$ is idempotent, that is $\mathbf{M^0}'\mathbf{M^0} = \mathbf{M^0}$

Exercise

Show that in a regression with only a constant, that is, when $\mathbf{X} = \mathbf{i}$, $\hat{\beta}$ is the mean of $\mathbf{Y}$ and the residual $\hat{\mathbf{e}}$ is $\mathbf{Y}$'s deviations from the mean.

The residual maker

Define

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'$$

Then

$$\hat{\mathbf{e}} = \mathbf{Y} - \mathbf{X}\hat{\beta} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}')\mathbf{Y} = \mathbf{MY}$$

M is called the ”residual maker”. Show that it is symmetric and idempotent.

Partitioned regression

Suppose the regression involves two sets of variables:

$$\mathbf{X} = [\mathbf{X_1} \quad \mathbf{X_2}].$$

So that

$$\mathbf{Y} = \mathbf{X_1}\boldsymbol{\beta_1} + \mathbf{X_2}\boldsymbol{\beta_2} + \mathbf{e}$$

Define

$$\mathbf{M_1} = \mathbf{I} - \mathbf{X_1}(\mathbf{X_1'}\mathbf{X_1})^{-1}\mathbf{X_1'}$$

$$\mathbf{M_2} = \mathbf{I} - \mathbf{X_2}(\mathbf{X_2'}\mathbf{X_2})^{-1}\mathbf{X_2'}$$

Write out the expression for $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ separately.