Intermediate Econometrics

Linear Algebra

Jean-Marc Robin

Sciences-Po

January 17, 2013

Vectors and vector spaces

- A vector is a set of numbers. Like $\mathbf{x} = (x_1, x_2, ..., x_N)$.
- A vector space is a set of vectors such that the sum of two vectors and the multiplication of a vector by a number make sense:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ...)$$
$$\lambda \mathbf{x} = (\lambda x_1, \lambda x_2, ...)$$

- A scalar product is an operation on two vectors onto $\mathbb{R}$:

$$\mathbf{x} \bullet \mathbf{y} = x_1 y_1 + x_2 y_2 + ...$$

  - $\mathbf{x} \bullet \mathbf{y} > 0$ if the vectors point in the same direction, $= 0$ if they are orthogonal.

Matrices as operators on vectors

- A matrix is a table of numbers, e.g. $A = \begin{pmatrix} 1.1 & 4 & -.7 \\ .5 & -3 & .9 \end{pmatrix} \in \mathbb{R}^{2\times 3}$.
- A matrix $A \in \mathbb{R}^{n\times m}$ operates on a vector $\mathbf{x} \in \mathbb{R}^m$ to produce another vector $\mathbf{y} \in \mathbb{R}^n$, the elements of which are linear combinations of the elements of $\mathbf{x}$.
- Let $m = 2, n = 3$,

$$\mathbf{y} = \underset{(2\times 3)}{A}\mathbf{x}$$

means that

$$\begin{aligned}
\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} x_2 + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}
\end{aligned}$$

Matrix transpose

- For any $A = [a_{ij}] \in \mathbb{R}^{n \times m}$ there exists a matrix $A^{\mathsf{T}} = [a_{ji}] \in \mathbb{R}^{m \times n}$ such that, for all $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$,

$$(A\mathbf{x}) \bullet \mathbf{y} = \mathbf{x} \bullet (A^{\mathsf{T}}\mathbf{y})$$

- This matrix takes a very simple form: if $A = [a_{ij}]$ has generic element a_{ij} (i is the row index and j is the column index) then $A^{\mathsf{T}} = [a_{ji}]$, i.e. writes the columns of A as the rows of A^{T}.

Vectors as matrices

- A single-column matrix $A \in \mathbb{R}^{n\times 1}$ operates on scalars to produces vectors of dimension n. In particular $A = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ applied to 1 produces a vector with coordinates a_{11} and a_{12}.
- A vector can be represented as a matrix in two ways: for example, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^{2\times 1}$ is a column-vector and $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{1\times 2}$ is a row-vector.

Matrix addition

- What $A\mathbf{x} + B\mathbf{x}$?
- Let $A = [a_{ij}]$ and $B = [b_{ij}]$. Then,

$$A\mathbf{x} + B\mathbf{x} = C\mathbf{x}$$

with for

$$C = [a_{ij} + b_{ij}].$$

Matrix multiplication

- Let $\underset{(n\times 1)}{\mathbf{y}} = \underset{(n\times m)}{A}\underset{(m\times 1)}{\mathbf{x}}$.
- What is $\underset{(p\times n)}{B}\mathbf{y}$?

$$B\mathbf{y} = B(A\mathbf{x}) = C\mathbf{x}$$

with

$$\underset{p\times m}{C} = [c_{ij}] = \underset{(p\times n)}{B}\underset{(n\times m)}{A}$$

where

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} b_{ik}a_{kj} \\ &= b_{i1}a_{1j} + .. + b_{in}a_{nj} \\ &= i\text{th row of } B \times j\text{th column of } A \end{aligned}$$

for $i = 1, ..., p, j = 1, ..., m$.

  - Matrices must be **conformable:** The number of columns of B must be equal to the number of rows of A.

Matrix transpose (again)

- The scalar product of two vectors becomes the following matrix multiplication. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n\times 1}$ be two column vectors of dimension n,

$$\underbrace{\mathbf{x} \bullet \mathbf{y}}_{\text{scalar product}} = \underbrace{\mathbf{x}^{\mathsf{T}}\mathbf{y}}_{\text{matrix multiplication}}$$

- Transpose of a product of matrices:

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$$

  Attention to the ordering change.
  - So:

$$\mathbf{y} = A\mathbf{x} \Leftrightarrow \mathbf{y}^{\mathsf{T}} = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}$$

  - and

$$(A\mathbf{x}) \bullet \mathbf{y} = (A\mathbf{x})^{\mathsf{T}}\mathbf{y} = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}\mathbf{y} = \mathbf{x} \bullet (A^{\mathsf{T}}\mathbf{y})$$

Matrix rank and determinant

- Let $A \in \mathbb{R}^{m \times n}$. The rows (resp. columns) of A are linearly independent if they cannot be written as a linear combination of the other rows (resp. columns).
- If the rows (resp. columns) of A are linearly independent then A has **full row (resp. column) rank** (by opposition to **rank defficient**).
- If $m > n$ then A is necessarily row rank defficient. If $m < m$ it is necessarily column rank defficient.
- The rank of a matrix is the maximal number of linearly independent rows or columns of A.
- It holds that
$$\text{rk}(A) = \text{rk}(A^{\mathsf{T}}) = \text{rk}(A^{\mathsf{T}}A) = \text{rk}(AA^{\mathsf{T}})$$
- If A is square, i.e $m = n$, then A has full rank if and only if its determinant is non nil:
$$\text{rk}(A) = n \Leftrightarrow \det(A) \neq 0$$
Otherwise, it is said **singular**.

Matrix inverse

- Consider a system of m linear equations with n unknowns $\mathbf{x} = (x_1, ..., x_n)^\mathsf{T} \in \mathbb{R}^{n\times 1}$:

$$\underset{(m\times 1)}{\mathbf{y}} = \underset{(m\times n)}{A}\underset{(n\times 1)}{\mathbf{x}}$$

  - If $m < n$ (less equations than unknown) the system has an infinite number of solutions;
  - If $m > n$ (strictly more equations), the system has no solution unless $m - n$ equations are redundant (i.e. $=$ linear combinations of n other ones);
  - If $m = n$ and A has full rank, then there exists a unique solution that is a vector of linear combinations of $\mathbf{y}$, i.e. there exists a matrix $A^{-1} \in \mathbb{R}^{n\times m}$ such that

$$\mathbf{x} = A^{-1}\mathbf{y}.$$

    This matrix is called the inverse of A.

- For example, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $\det(A) = ad - bc$ and $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$. Formulas exists for higher dimensions.

Trace

- $\mathrm{tr}A$ is the sum of the diagonal elements of A
- It is equal to the sum of the characteristic roots.
- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

Cholesky decomposition

- If A has real entries and is symmetric and positive definite, then A can be decomposed as $A = LL^{\mathsf{T}}$, where L is a lower triangular matrix with strictly positive diagonal entries.

Characteristic roots

- Roots of the determinental equation:

$$|A - \lambda I| = 0$$

- Polynomial in the dimension of A.
- The characteristic roots are the eigenvalues of A.
- They are in general complex. They are real for a symmetric matrix.
- Let $\lambda_1 > \lambda_2 > ... > \lambda_n$ be the char. roots of a symmetric matrix A. Then

$$\lambda_1 = \max_x \frac{x^\top Ax}{x^\top x} = \frac{x_1^\top Ax_1}{x_1^\top x_1}$$

$$\lambda_2 = \max_{x^\top x_1 = 0} \frac{x^\top Ax}{x^\top x} = \frac{x_2^\top Ax_2}{x_2^\top x_2}$$

$$\lambda_3 = \max_{\substack{x^\top x_1 = 0 \\ x^\top x_2 = 0}} \frac{x^\top Ax}{x^\top x} = \frac{x_3^\top Ax_3}{x_3^\top x_3}$$

...

Diagonalization of general matrices

- For any square matrix A, with distinct characteristic roots, there exists a nonsingular matrix P such that $PAP^{-1} = \Lambda$, where Λ is the diagonal matrix with the characteristic roots in the diagonal and the columns of which are eigenvectors of A.
- If the char. roots are not distincts, A takes the Jordan canonical form, ie $A = P^{-1}JA$, where J is nearly diagonal. For example, with $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = \lambda_4 = \lambda_5 = 4$,

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Diagonalization of symmetric matrices

- Every real symmetric matrix A can be diagonalized, moreover the eigen decomposition takes a simpler form:

$$A = Q\Lambda Q^{\mathsf{T}}$$

where Q is an orthogonal matrix (the columns of which are eigenvectors of A), and Λ is real and diagonal (having the eigenvalues of A on the diagonal).

Positivity

- For a symmetric matrix A, A is positive definite (write $A \succ 0$) if
  - $x^{\mathsf{T}} A x > 0$ for all $x \neq 0$
  - or equivalently if its characteristic roots are all positive
- Semi positive definite ($A \succeq 0$) if $x^{\mathsf{T}} A x \geq 0$ for all $x \neq 0$
- Let A, B be symmetrix, nonsingular matrices of same size. Then $A \succeq B \succeq 0$ implies $B^{-1} \geqq A^{-1}$.

Projection

- Let $A \in \mathbb{R}^{n \times r}$ be a matrix of rank r. A matrix of the form $P = A(A^\mathsf{T} A)^{-1} A^\mathsf{T}$ is called a projection matrix.
- P is symmetrix and idempotent: $P = P^\mathsf{T} = P^2$
- $\mathrm{rk}(P) = n - r$ and the eigenvalues of P are 1 with multiplicity $n - r$ and 0 with multiplicity r.
- If $x = Ac$ for some c then $Px = x$ (hence projection)
- $M = I - P$ is also a projection matrix and $Mx = 0$ for all $x = Ac$

Block-matrices

- For example, $C = (A, B)$ or $E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.
- The same rules defining the addition and the multiplication for matrices apply to block-matrices, if blocks are conformable. For example,

$$(A, B)\begin{pmatrix} C \\ D \end{pmatrix} = AC + BD$$

$$(A, B)^{\mathsf{T}} = \begin{pmatrix} A^{\mathsf{T}} \\ B^{\mathsf{T}} \end{pmatrix}$$

$$\begin{pmatrix} C \\ D \end{pmatrix}^{\mathsf{T}} = (C^{\mathsf{T}}, D^{\mathsf{T}})$$

- $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \left|A - BD^{-1}C\right|$ if $|D| \neq 0$
- $\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} E^{-1} & -E^{-1}BD^{-1} \\ -D^{-1}CE^{-1} & F^{-1} \end{pmatrix}$ with $E = A - BD^{-1}C, F = D - CA^{-1}B$

Kronecker product

- Let $A = (a_{ij})$ and B be matrices. Then

$$A \otimes B = (a_{ij}B) = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1L}B \\ a_{21}B & a_{22}B & \cdots & a_{2L}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1}B & a_{M2}B & \cdots & a_{ML}B \end{pmatrix}.$$

- Properties:

$$\begin{aligned} (A \otimes B)^{\mathsf{T}} &= (A^{\mathsf{T}}) \otimes (B^{\mathsf{T}}) \\ (A \otimes B)^{-1} &= (A^{-1}) \otimes (B^{-1}) \\ (A \otimes B)(C \otimes D) &= (AC) \otimes (BD) \end{aligned}$$

(if A and B are non singular and if A and C and B and D are conformable).

- $\mathrm{vec}(ABC) = (C^{\mathsf{T}} \otimes A)\mathrm{vec}(B)$