Economics 142

Quick Review of Vector and Matrix Notation

u, a vector, is a k-element list. We will assume that vectors are "column vectors" unless otherwise noted. So $u' = (u_1, u_2, ...u_k)$. If u, v are two k-vectors, then $u'v = \sum_i u_i v_i$, the "dot product."

A, a matrix, is an $l \times k$ "table" with (i, j) element a_{ij}, so:

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1k} \\ a_{21} & a_{22} & ... & a_{2k} \\ ... & & & ... \\ a_{l1} & a_{l2} & ... & a_{lk} \end{pmatrix}$$

If A is $l \times k$ and v is $k \times 1$ then

$$Av = \begin{pmatrix} \sum_j a_{1j} v_j \\ \sum_j a_{2j} v_j \\ ... \\ \sum_j a_{lj} v_j \end{pmatrix} = \begin{pmatrix} a'_{1.} v \\ a'_{2.} v \\ ... \\ a'_{l.} v \end{pmatrix},$$

where $a_{1.}$ means "the first row of A" (and similarly, $a_{.h}$ means the h^{th} column of A). A special matrix is I, the identity matrix, which has 1 on the diagonals and 0 off diagonal. Sometimes people try to be careful and write I_k to mean the $k \times k$ identity matrix. Note that for any vector u, $Iu = u$.

If A is $l \times k$ and B is $k \times m$ then

$$AB = \begin{pmatrix} a_{11} & ... & a_{1k} \\ a_{21} & ... & a_{2k} \\ ... & & ... \\ a_{l1} & ... & a_{lk} \end{pmatrix} \begin{pmatrix} b_{11} & ... & b_{1m} \\ b_{21} & ... & b_{2m} \\ ... & & ... \\ b_{l1} & ... & b_{lm} \end{pmatrix} = \begin{pmatrix} a'_{1.} b_{.1} & ... & a'_{1.} b_{.m} \\ a'_{2.} b_{.1} & ... & a'_{2.} b_{.m} \\ ... & & ... \\ a'_{l.} b_{.1} & ... & a'_{l.} b_{.m} \end{pmatrix}$$

The matrices have to be "conformable", so A has as many columns as B has rows, so that you can form the o, p element as the dot-product of the o^{th} row of A with the p^{th} column of B.

Note that for a "square" matrix A,$AI = A$, and $IA = A$, where I has the same size as A.

A common use of matrices is to represent systems of linear equations. For example, suppose you have k equations of the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1k}x_k &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2k}x_k &= b_2 \\ &... \\ a_{k1}x_1 + a_{k2}x_2 + ... + a_{kk}x_k &= b_k \end{aligned}$$

This system can be represented as the matrix equation $Ax = b$. A unique solution for x will exist if A has "full rank", which means that there are k linearly independent equations in the system: you cannot take weighted sums of any $k-1$ of the equations and get the kth one back. If this is true, then A is "invertible" – there exists another $k \times k$ matrix A^{-1},such that $A^{-1}A = AA^{-1} = I$. So the solution to the system $Ax = b$ is $x = A^{-1}b$. There are lots of ways of expressing the fact that A has full rank. One useful one is: "$Av = 0$ only holds for $v = 0$" i.e., the "null space" of A contains only the 0-vector. Another one is that $det(A) \neq 0$ where $det(.)$ means the determinant of A.

Variance-covariance and covariance matrices

Suppose we have observe a sample of size N on a set of covariates $X_1, X_2 ... X_k$ and some other variable of interest$Y$. For the i^{th}person in the sample we observe $x_i' = (x_{1i}, x_{2i,} x_{3i} ... x_{ki})$ and y_i. The *mean of* X is the vector

$$\bar{x} = \frac{1}{N}\sum_i x_i = \begin{pmatrix} \frac{1}{N}\sum_i x_{1i} \\ \frac{1}{N}\sum_i x_{2i} \\ ... \\ \frac{1}{N}\sum_i x_{ki} \end{pmatrix}$$

The *variance-covariance matrix* of the $X's$ is the matrix:

$$V_{xx} = \frac{1}{N}\sum_i \begin{pmatrix} (x_{1i}-\bar{x}_1)^2 & (x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2) & ... & (x_{1i}-\bar{x}_1)(x_{ki}-\bar{x}_k) \\ (x_{2i}-\bar{x}_2)(x_{1i}-\bar{x}_1) & (x_{2i}-\bar{x}_2)^2 & & (x_{2i}-\bar{x}_2)(x_{ki}-\bar{x}_k) \\ ... & & & \\ (x_{ki}-\bar{x}_k)(x_{1i}-\bar{x}_1) & (x_{ki}-\bar{x}_k)(x_{2i}-\bar{x}_2) & ... & (x_{ki}-\bar{x}_k)^2 \end{pmatrix}$$

Note that if we call

$$\hat{s}_{jk}^2 = \frac{1}{N}\sum_i (x_{ji}-\bar{x}_j)(x_{ki}-\bar{x}_k)$$

then V_{XX} is just the matrix of $\hat{s}_{jk}^2$ terms. The diagonal elements are the estimated variances. (Again, note we are using N instead of $N-1$ for degrees of freedom). The *covariance of* X *with* Y is the vector:

$$C_{xy} = \frac{1}{N}\sum_i \begin{pmatrix} (x_{1i}-\bar{x}_1)(y-\bar{y}) \\ (x_{2i}-\bar{x}_2)(y-\bar{y}) \\ ... \\ (x_{ki}-\bar{x}_k)(y-\bar{y}) \end{pmatrix}.$$

A useful fact to note is that

$$\begin{aligned} \hat{s}_{jk}^2 &= \frac{1}{N}\sum_i (x_{ji}-\bar{x}_j)(x_{ki}-\bar{x}_k) \\ &= \frac{1}{N}\sum_i (x_{ji}-\bar{x}_j)x_{ki} \\ &= \frac{1}{N}\sum_i x_{ji}^2 - \bar{x}_j^2 \end{aligned}$$

So you can compute the var-cov matrix in several different ways. Lots of times people convert covariances into correlations. For any 2 of the X's:

$$\hat{\rho}_{jk} = \frac{\hat{s}^2_{jk}}{\sqrt{\hat{s}^2_{kk}\hat{s}^2_{jj}}}$$

The correlation has to fall between -1 and +1. If the correlation of two $X's$ is 1, then they are related as $x_{ji} = a + bx_{ki}$ with no "residual" (where $b > 0$).