smfl11.tex
Lecture 11. 11.2.2011

III. MULTIVARIATE ANALYSIS

1. PRELIMINARIES: MATRIX THEORY.

Modern Algebra splits into two main parts: Groups, Rings and Fields on the one hand, and Linear Algebra on the other. Linear Algebra deals with *linear transformations* between *vector spaces*. We confine attention here to the *finite-dimensional* case; the infinite-dimensional case needs Functional Analysis and is harder. Broadly, Parametric Statistics can be handled in finitely many dimensions, Non-Parametric Statistics needs infinitely many.

Given a finite-dimensional vector space V, we can always choose a *basis* (a maximal set of linearly independent vectors). All such bases contain the same number of vectors; if this is n, the vector space has *dimension* n.

Given two finite-dimensional vector spaces and a linear transformation α between the two, choice of bases $(e_1, \ldots, e_m)$ and $(f_1, \ldots, f_n)$ determines a *matrix* $A = (a_{ij})$ by

$$e_i \alpha = \sum_{j=1}^{n} a_{ij} f_j \qquad (i = 1, \ldots, m).$$

We write

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix},$$

or $A = (a_{ij})$ more briefly. The a_{ij} are called the *elements* of the matrix; we write A $(m \times n)$ for A (m rows, n columns).

Matrices may be subjected to various operations:

1. *Matrix addition.* If $A = (a_{ij})$, $B = (b_{ij})$ have the same size, then

$$A \pm B := (a_{ij} \pm b_{ij})$$

(this represents $\alpha \pm \beta$ if α, β are the underlying linear transformations).

2. *Scalar multiplication.* If $A = (a_{ij})$ and c is a scalar (real, unless we specify complex), then the matrix

$$cA := (ca_{ij})$$

represents $c\alpha$.
3. *Matrix multiplication.* If A is $m \times n$, B is $n \times p$, then $C := AB$ is $n \times p$, where $C = (c_{ij})$ and

$$c_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj}$$

(this represents the product, or composition, $\alpha\beta$ or $x \mapsto x\alpha\beta$).
Note. Matrix multiplication is non-commutative! – $AB \neq BA$ in general, even when both are defined (which can only happen for A, B square of the same size).
Partitioning.

We may *partition* a matrix A in various ways. for instance, A as above partitions as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is $r \times s$, A_{12} is $r \times (n-s)$, A_{21} is $(m-r) \times s$, A_{22} is $(m-r) \times (n-s)$, etc. In the same way, A may be partitioned as
(i) a column of its rows; (ii) a row of its columns.
Rank.

The maximal number of linearly independent *rows* of A is always the same as the maximal number of independent *columns*. This number, r, is called the *rank* of A. When $r = \min(m, n)$ is as big as it could be, the matrix A has *full rank*.
Inverses.

When a square matrix A $(n \times n)$ has full rank n, the linear transformation $\alpha : V \to V$ that it represents is *invertible*, and so has an inverse map $\alpha^{-1} : V \to V$ such that $\alpha\alpha^{-1} = \alpha^{-1}\alpha = i$, the identity map, and α^{-1} is also a linear transformation. The matrix representing α^{-1} is called A^{-1}, the *inverse matrix* of A:

$$AA^{-1} = A^{-1}A = I,$$

the *identity matrix* of size n: $I = (\delta_{ij})$ ($\delta_{ij} = 1$ if $i = j$, 0 otherwise – the *Kronecker delta*).
Transpose.

If $A = (a_{ij})$, the *transpose* is A', or $A^T := (a_{ji})$.

Note that, when all the matrices are defined,

$$(AB)^{-1} = B^{-1}A^{-1}$$

(as this gives $(AB)(AB)^{-1} = ABB^{-1}A^{-1} = AA^{-1} = I$, and similarly $(AB)^{-1}(AB) = I$, as required), and

$$(AB)^T = B^T A^T$$

(as this has (i,j) element $\sum_k (B^T)_{ik}(A^T)_{kj} = \sum_k b_{ki}a_{jk} = \sum_k a_{jk}b_{ki} = (AB)_{ji}$, as required).

Determinants.

There are $n!$ *permutations* σ of the set

$$N_n := \{1, 2, \ldots, n\}$$

– bijections $\sigma : N_n \to N_n$. Each permutation may be decomposed into a product of *transpositions* (interchanges of two elements), and the *parity* of the number of transpositions in any such decomposition is always the same. Call σ *odd* or *even* according as this number is odd or even. Write

$$\text{sgn } \sigma := 1 \quad \text{if } \sigma \text{ is even}, \quad -1 \quad \text{if } \sigma \text{ is odd}$$

for the *sign* or *signum* of σ. For A a square matrix of size n, the function

$$\det A, \text{ or } |A|, := \sum_{\sigma} (-1)^{\text{ sgn } \sigma} a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)},$$

where the summation extends over all $n!$ permutations, is called the *determinant* of A, $\det A$ or $|A|$.

Properties.

1. $|A^T| = |A|$.

Proof. If σ^{-1} is the inverse permutation to σ, σ and σ^{-1} have the same parity, so the sums for their determinants have the same terms, but in a different order.

2. If two rows (or columns) of A coincide, $|A| = 0$.

Proof. Interchanging two rows changes the sign of $|A|$ (extra transposition, which changes the parity), but leaves A and so $|A|$ unaltered (as the two rows coincide). So $|A| = -|A|$, giving $|A| = 0$.

3. $|A|$ depends linearly on each row (or column) (det is a *multilinear* function, and this area is called Multilinear Algebra).

4. If A is $n \times n$, $|A| = 0$ iff A has rank $r < n$.

5. *Multiplication Theorem for Determinants.* If A, B are $n \times n$ (so AB, and BA, are defined),

$$|AB| = |A|.|B|.$$

6. *Inverses again.*

If A is $n \times n$, the (i, j) *minor* is the determinant A_{ij} of the $(n-1) \times (n-1)$ submatrix obtained by deleting the ith row and jth column. The (i, j) *cofactor*, or *signed minor*, is $(-)^{i+j} A_{ij}$ (the signs follow a chessboard or chequerboard pattern, with $+$ in the top left-hand corner),

The matrix $B = (b_{ij})$, where

$$b_{ij} := (-)^{i+j} A_{ji}/|A|,$$

ia the *inverse matrix* A^{-1} of A, defined iff $|A| \neq 0$ (A is called *singular* if $|A| = 0$, *non-singular* otherwise (thus a square matrix has a determinant iff it is non-singular), and

$$AA^{-1} = A^{-1}A = I$$

as before:

inverse = transposed matrix of cofactors over determinant.

Proof. With B as above, $C := AB = (c_{ij})$,

$$c_{ij} = \sum_k a_{ik} b_{kj} = \sum_k a_{ik}.(-)^{k+j} A_{jk}/|A|.$$

If $i = j$, the RHS is 1 (expansion of $|A|$ by its ith row). If not, the RHS is 0 (expansion of the determinant of a matrix with two identical rows). So $c_{ij} = \delta_{ij}$, so $C = AB = I$. Similarly, $BA = I$.

Solution of linear equations.

If A is $n \times n$, the linear equations

$$Ax = b$$

possess a unique solution x iff A is non-singular (A^{-1} exists), and then

$$x = A^{-1}b.$$

If A is singular (A has rank $r < n$), then *either* there is *no* solution (the equations are *inconsistent*), *or* there are *infinitely many* solutions (some equations are *redundant*, and one can give some of the elements x_i *arbitrary* values and solve for the rest). What decides between these two cases is the rank of the *augmented* matrix (A, b) obtained by adjoining the vector b as a final column.