Holger Broman, August 19, 1999

Determinants

The story about determinants is messy. The definition is awkward, most formulas are boring in their details; some are, however, very useful. This leaflet will therefore end with the definition, and start with simple examples and some unproven properties of determinants. One particularly important property of determinants will be traced, namely the relation between determinants and hypervolumes.

The start is by 2×2 matrices. By the way, the concept of determinants of matrices applies only to square matrices. Here is a definition applicable to the 2×2 case:

$$\det\, A = |A| = \left| \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right| = ad - bc$$

Note 1: It is generally true that the determinant of an $N \times N$ matrix is a mapping from $\mathbb{R}^{N\times N}$ to $\mathbb{R}^1$. Use $\mathbb{C}$ if needed.

Note 2: det $A = 0$ iff $ad = bc$, which makes the column vectors of A linearly dependent. This fact also extrapolates to the $N \times N$ case.

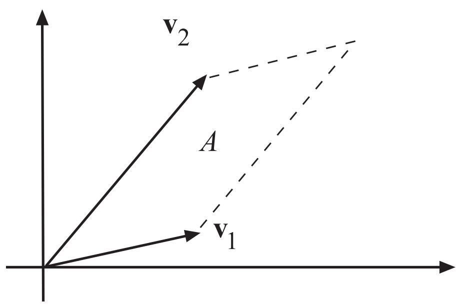
Note 3: det $A = 0$ iff at least one eigenvalue of A equals zero. This fact also.......

Exercise: Show the statement of Note 3 by solving the equation

$$\det\,(A - \lambda\, I) = 0.$$

Note 4: det $A = 0$ iff A is not invertible. This follows directly from Note 2, and also extrapolates......

Armed with these properties, we can start our journey to find the relation between determinants and hypervolumes. Let us start with a simple case:



"Seek the area A spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$". The area of a parallelepiped is the length of the base times the length of the height. Choose the base to be $\mathbf{v}_1$, the length is $|\mathbf{v}_1|$. The height then becomes the part of $\mathbf{v}_2$ perpendicular to $\mathbf{v}_1$, and we know all about that from the leaflet on projection matrices:

$$A = |\mathbf{v}_1| \left|P^{\perp}\mathbf{v}_2\right| = |\mathbf{v_1}| \left|(I - P)\mathbf{v}_2\right|,$$

where P is the projection matrix that projects onto $\mathbf{v}_1$:

$$P = \mathbf{v}_1(\mathbf{v}_1^T\,\mathbf{v})^{-1}\mathbf{v}^T = |\mathbf{v}_1|^{-2}\,\mathbf{v}_1\mathbf{v}_1^T$$

Let

$$\mathbf{v}_1^T = (a, b), \quad \mathbf{v}_2^T = (c, d)$$

and let us get rid of the square-roots

$$A^2 = |\mathbf{v}_1|^2 \left|P^{\perp}\,\mathbf{v}_2\right|^2.$$

$$A^2 = \mathbf{v}_1^T\mathbf{v}_1\ \mathbf{v}_2^T(P^{\perp})^T(P^{\perp})\mathbf{v}_2 = \ (P^{\perp}\text{ symmetric}) =$$

$$= \mathbf{v}_1^T\mathbf{v}_1\ \mathbf{v}_2^T(P^{\perp})^2\mathbf{v}_2 = (P^{\perp}\text{ is a projection matrix}) =$$

$$= \mathbf{v}_1^T\mathbf{v}_1\ \mathbf{v}_2^T P^{\perp}\mathbf{v}_2 =$$

$$= \text{some manipulations, please do them} =$$

$$= (a^2 + b^2)\ \frac{1}{a^2 + b^2}(ad - bc)^2 = (ad - bc)^2$$

How extraordinarily nice:

$$A^2 = \left\{\det[\mathbf{v}_1\ \mathbf{v}_2]\right\}^2$$

Do you think that this will extrapolate to N vectors in N dimensions?

Do you believe that the following formula for the hypervolume, H_M, spanned by M vectors in N dimensions is correct?

$$B = [\mathbf{v}_1 \ldots \mathbf{v}_M] \ \text{ – not square}$$

$$H_M^2 = \det\left[B^T B\right]?$$

If your answer is yes to both questions, you get full score. Before we proceed to prove that, we will make a tour into an important application.

Substitution of variables in two-dimensional integrals

Given the following

$$\iint\limits_D f(x, y)\, dx\, dy,$$

we decide that it will be simpler to solve the problem in the $u - v$ plane

$$\begin{cases} u = u(x, y) \\ v = v(x, y) \end{cases}$$

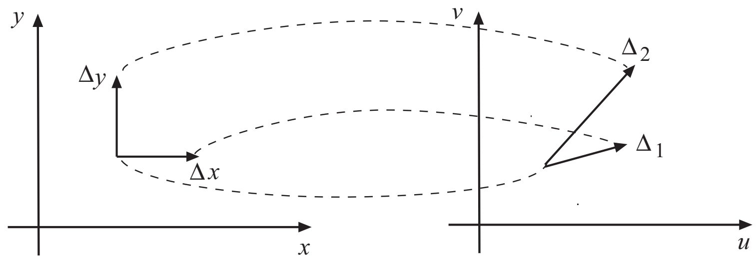
Note 5: In the general case, we would prefer the following notation for the substitution:

$$\mathbf{g}: \quad \mathbb{R}^N \ni \mathbf{x} \overset{\mathbf{g}}{\curvearrowright} \mathbf{y} \in \mathbb{R}^N.$$

i.e.

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) \quad \text{or} \quad \mathbf{y} = \mathbf{y}(\mathbf{x}).$$

Now, back to x, y and u, v. What happens to an area in the $x - y$ plane when transferred to the $u - v$ plane?



Mapping by $u = u(x, y)$, $v = v(x, y)$.

In the $x - y$ plane it is simple,

$$A_{xy} = \Delta x\ \Delta y$$

For us, it is also simple in the $u - v$ plane,

$$A_{uv} = \left|\det\left[\boldsymbol{\Delta}_1\ \boldsymbol{\Delta}_2\right]\right|,$$

so let us calculate $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$:

$$\boldsymbol{\Delta}_1 = \begin{bmatrix} u(x + \Delta x, y) - u(x, y) \\ v(x + \Delta x, y) - v(x, y) \end{bmatrix} \longrightarrow \begin{bmatrix} \Delta x\ u'_x \\ \Delta x\ v'_x \end{bmatrix} \text{ as } \Delta x \to 0$$

$$\boldsymbol{\Delta}_2 = \cdots \longrightarrow \begin{bmatrix} \Delta y\ u'_y \\ \Delta y\ v'_y \end{bmatrix} \text{ as } \Delta y \to 0$$

given that the mapping $(x, y) \curvearrowright (u, v)$ is "nice". So,

$$A_{uv} \longrightarrow \left\| \begin{matrix} \Delta x\ u'_x & \Delta y\ u'_y \\ \Delta x\ v'_x & \Delta y\ v'_y \end{matrix} \right\| = \Delta x\ \Delta y \left| \det \begin{bmatrix} u'_x\ u'_y \\ v'_x\ v'_y \end{bmatrix} \right|$$

Note 6: The matrix

$$\begin{bmatrix} u'_x & u'_y \\ v'_x & v'_y \end{bmatrix}$$

is called the Jacobian (matrix) of the mapping $(x, y) \curvearrowright (u, v)$.

The important conclusion is:

> The area scaling of a substitution of variables in two dimensions is given by the absolute value of the determinant of the Jacobian.

Do you think this will extrapolate to N-dimensional substitutions?

—————— o O o ——————

Hypervolumes

Given M vectors in $\mathbb{R}^N$ $(M \leq N)$ we want to know the hypervolume, H_M, of the parallelepiped spanned by those vectors. If all vectors are orthogonal, there is little doubt regarding the proper definition. So let us make all things perpendicular.

$$\begin{aligned} H_1^2 &= |\mathbf{v}_1|^2 & & \\ H_2^2 &= |\mathbf{v}_1|^2 \left|P_1^\perp \mathbf{v}_2\right|^2 & \qquad & P_1 \text{ projects onto } \mathbf{v}_1 \\ H_3^2 &= H_2^2 \left|P_2^\perp \mathbf{v}_3\right|^2 & \qquad & P_2 \text{ projects onto the } \mathbf{v}_1, \mathbf{v}_2 \text{ plane} \end{aligned}$$

We arrive at

$$H_M^2 = H_{M-1}^2 \left|P_{M-1}^\perp \mathbf{v}_M\right|^2 = \prod_{m=1}^{M} \left|P_{m-1}^\perp \mathbf{v}_m\right|^2$$

if we let

$$P_0^\perp \triangleq I.$$

Now we need to prove that this equals

$$H_M^2 = \det[B^T B],$$

where

$$B = [\mathbf{v}_1 \cdots \mathbf{v}_M]$$

Note 7: $B^T B$ is of dimensions $M \times M$, so its determinant will contain $M!$ terms, each with M factors. This property follows from the definition still to come. We would very much like things to be simpler.

There is a guide on how to make things simple in noting that if the vectors $\mathbf{v}_m$ were orthogonal, then B would have orthogonal columns and $B^T B$ would have been a diagonal matrix. The determinant of a diagonal matrix has only one term, namely the product of the diagonal entries. This fact actually holds also for triangular matrices. If you are familiar with QR factorizations – described in another leaflet – the road to follow is clear as crystal: find matrices Q and R so that

$$B = QR,$$

Q has orthonormal columns and R is square, triangular. Then

$$\det(B^T B) = \det(R^T\, Q^T\, Q\, R) = \det(R^T\, I\, R) = (\det R)^2.$$

We have used two properties of determinants in the above argument:

$$\begin{aligned} \det R^T &= \det R, \\ \det AB &= \det A \cdot \det B, \end{aligned}$$

still to be claimed. This means that the expression for H_M^2 would have only one term.

Let us start. The columns of B span a space, and we want the columns of Q to be an orthogonal basis of the same space:

$$B = (\mathbf{v}_1 \cdots \mathbf{v}_M) = QR = \left(\mathbf{v}_1\, P_1^\perp\, \mathbf{v}_2 \cdots P_{M-1}^\perp\, \mathbf{v}_M\right) \cdot R$$

Note 8: In this choice, Q is not orthonormal, but has orthogonal columns.

Remains to find R. Now recall the interpretation of a matrix-vector product as a linear combination of the columns of the matrix. This gives the first column of R:

$$(\mathbf{v}_1 \cdots \mathbf{v}_M) = \left(\mathbf{v}_1\, P_1^\perp\, \mathbf{v}_2 \cdots P_{M-1}^\perp\, \mathbf{v}_M\right) \begin{pmatrix} 1 & \\ 0 & \text{rest} \\ \vdots & \text{of} \\ \vdots & \text{R} \\ 0 & \end{pmatrix},$$

as the $\mathbf{v}_1$ in Q must map directly on the $\mathbf{v}_1$ in B, no $\mathbf{v}_2$:s etc. are allowed. Now for the second column. $P_1^\perp\, \mathbf{v}_2$ equals $\mathbf{v}_2$ — $P_1\, \mathbf{v}_2$ equals $\mathbf{v}_2$ minus the part of $\mathbf{v}_2$ pointing in the $\mathbf{v}_1$ direction. We must thus add back that part to restore $\mathbf{v}_2$:

$$(\mathbf{v}_1\, \mathbf{v}_2 \cdots \mathbf{v}_M) = \left(\mathbf{v}_1\, P_1^\perp\, \mathbf{v}_2 \cdots P_{M-1}^\perp\, \mathbf{v}_M\right) \begin{pmatrix} 1 & x & \\ 0 & 1 & \text{rest} \\ \vdots & 0 & \text{of} \\ \vdots & \vdots & \text{R} \\ 0 & 0 & \end{pmatrix},$$

Now, everything falls into place:

$$R = \begin{bmatrix} 1 & x & \cdots & \cdots & x \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & x \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix},$$

$$\det R = 1$$

$$Q = \left[\mathbf{v}_1 \; P_1^{\perp} \mathbf{v}_1 \cdots P_{M-1}^{\perp} \mathbf{v}_M \right]$$

$$\det B^T B = \det \left(R^T Q^T Q R \right) = \det R^T \det \left(Q^T Q \right) \det R =$$

$$= \det \begin{bmatrix} \mathbf{v}_1^T \mathbf{v}_1 & & 0 \\ & \mathbf{v}_2^T P_1^{\perp} \mathbf{v}_2 & \\ 0 & & \ddots \\ & & & \mathbf{v}_M^T P_{M-1}^T \mathbf{v}_M \end{bmatrix} = H_M^2$$

which proves the proposition.

——————— o O o ———————

Substitution of variables in N-dimensional integrals

This is now simple. The problem is

$$\int f(\mathbf{x})\, d\mathbf{x}$$

with the substitution

$$\mathbf{y} = \mathbf{y}(\mathbf{x}).$$

Introduce the axes-parallel displacements in the $\mathbf{x}$-space: $\Delta x_1 \cdots \Delta x_N$. As the displacements are orthogonal, the hypervolume is given by

$$H_{\mathbf{x}}^2 = \prod_{n=1}^{N} (\Delta x_n)^2$$

The displacement vector in the $\mathbf{y}$-space corresponding to displacement Δx_n is

$$\Delta y_n = \mathbf{y}(\mathbf{x} + \begin{bmatrix} 0 \\ \vdots \\ \Delta x_n \\ \vdots \\ 0 \end{bmatrix}) - y(\mathbf{x}) \longrightarrow \Delta x_n \begin{bmatrix} \frac{\partial y_1}{\partial x_n} \\ \vdots \\ \vdots \\ \frac{\partial y_N}{\partial x_n} \end{bmatrix} \quad \text{as } \Delta x_n \to 0,$$

so that

$$B = {}^{1} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_N} \\ \vdots & & \vdots \\ \frac{\partial y_N}{\partial x_1} & \cdots & \frac{\partial y_N}{\partial x_N} \end{bmatrix} \prod_{n=1}^{N} \Delta x_n = J \cdot \prod_{n=1}^{N} \Delta x_n$$

where J is the Jacobian of the substitution[2] $\mathbf{y} = \mathbf{y}(\mathbf{x})$. Thus the hypervolume scaling is

$$\frac{H_{\mathbf{y}}}{H_{\mathbf{x}}} = |\det\ J|\,.$$

Determinants, definition of

Even and odd permutations are defined by the number shifts of the positions needed to reach the new ordering of some elements. Examples follow:

$$(\ a\ \ b\ \ c\ ) \longrightarrow (\ b\ \ a\ \ c\ ) \qquad \text{odd}$$

$$(\ a\ \ b\ \ c\ ) \longrightarrow (\ c\ \ a\ \ b\ ) \qquad \text{even}$$

Let us associate a sign, $+1$ to even and -1 to odd permutations. The determinant of an N by N matrix is defined by the sum of the $N!$ terms you obtain by choosing one and only one element from each row and column and multiply them, including the sign of the permutation of rows (or columns, same thing):

$$\det\ A = \sum_{n=1}^{N!} (a_{1\alpha} \cdot a_{2\beta} \ldots \cdot a_{N\omega})\,\mathrm{sign}\,(\alpha\ \beta \ldots \omega)\,,$$

where the sum extends over all $N!$ permutations $(\alpha\ \beta \ldots \omega)$ of the numbers $(1 \ldots N)$.

An alternative way to define determinants is sketched here. As we know what the determinant of a 2×2 matrix is, we can use it to define the determinant of a 3×3 matrix:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =$$

$$= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Note 9: You see the pattern and how it can be extended to $N \times N$ matrices recursively.

[1] We have used a third unproven property of determinants here, namely that you can move a factor common to one row (column) of the matrix outside the expression for the determinant.

[2] You have certainly noticed the equality between the words "substitution" and "mapping".

Note 10: The alternation of the sign corresponds just to the even and odd permutations introduced above.

Note 11: The determinant of the matrix you obtain by deleting one row and one column is called a cofactor of the original matrix.

General properties – and proofs – can be found in any textbook on linear algebra.