planetmath.org

Math for the people, by the people.

normal equations

Canonical name	NormalEquations
Date of creation	2013-03-22 12:04:28
Last modified on	2013-03-22 12:04:28
Owner	akrowne (2)
Last modified by	akrowne (2)
Numerical id	7
Author	akrowne (2)
Entry type	Definition
Classification	msc 65-00
Classification	msc 15-00

Normal Equations

We consider the problem $Ax \approx b$, where A is an $m \times n$ matrix with $m \geq n$ rank $(A) = n$, b is an $m \times 1$ vector, and x is the $n \times 1$ vector to be determined.

The sign $\approx$ stands for the least squares approximation, i.e. a minimization of the norm of the residual $r = Ax - b$.

$$||Ax - b||_2 = ||r||_2 = \left[\sum_{i=1}^{m} r_i^2 \right]^{1/2}$$

or the square

$$\begin{aligned} F(x) &= \frac{1}{2}||Ax - b||_2^2 = \frac{1}{2}(Ax - b)^T(Ax - b) \\ &= \frac{1}{2}(x^T A^T Ax - 2x^T A^T b + b^T b) \end{aligned}$$

i.e. a differentiable function of x. The necessary condition for a minimum is:

$$\nabla F(x) = 0 \text{or} \frac{\partial F}{\partial x_i} = 0 (i = 1, \ldots, n)$$

These equations are called the normal equations , which become in our case:

$$A^T Ax = A^T b$$

The solution $x = (A^T A)^{-1} A^T b$ is usually computed with the following algorithm: First (the lower triangular portion of) the symmetric matrix $A^T A$ is computed, then its Cholesky decomposition LL^T. Thereafter one solves $Ly = A^T b$ for y and finally x is computed from $L^T x = y$.

Unfortunately $A^T A$ is often ill-conditioned and strongly influenced by roundoff errors (see [Golub89]). Other methods which do not compute $A^T A$ and solve $Ax \approx b$ directly are QR decomposition and singular value decomposition.

References

- Originally from The Data Analysis Briefbook (`http://rkb.home.cern.ch/rkb/titleA.html`