1 Linear least-squares problem

Linear least-squares problem is the problem of finding an approximate solution to an overdetermined system of linear equations. It arises often in applications where a theoretical model is fitted to experimental data.

Consider a linear system

$$A\mathbf{c} = \mathbf{b} \;, \tag{1}$$

where A is an $n \times m$ matrix, $\mathbf{c}$ is an m-component vector and $\mathbf{b}$ is an n-component vector. If the number of unknowns n is larger than the number of equations m, the system is called *overdetermined* and generally has no solution. However, it is still possible to find an approximate solution: the one which minimizes the Euclidean norm of the difference between $A\mathbf{c}$ and $\mathbf{b}$,

$$\min_{\mathbf{c}} \|A\mathbf{c} - \mathbf{b}\|^2 \;. \tag{2}$$

The problem (2) is called the linear least-squares problem and the vector $\mathbf{c}$ that minimizes $\|A\mathbf{c}-\mathbf{b}\|^2$ is called the *least-squares solution.*

The linear least-squares problem can be solved by QR-decomposition. The matrix A is factorized as $A = QR$, where Q is $n \times m$ matrix with orthogonal columns, $Q^TQ = 1$, and R is an $m \times m$ upper triangular matrix. The Euclidean norm

$$\begin{aligned} \|A\mathbf{c} - \mathbf{b}\|^2 &= \|QR\mathbf{c} - \mathbf{b}\|^2 \\ &= \|R\mathbf{c} - Q^T\mathbf{b}\|^2 + \|(1 - QQ^T)\mathbf{b}\|^2 \\ &\geq \|(1 - QQ^T)\mathbf{b}\|^2 \end{aligned} \tag{3}$$

can apparently be minimized by solving an $m \times m$ system of linear equations

$$R\mathbf{c} - Q^T\mathbf{b} = 0 \tag{4}$$

by back-substitution.

1.1 Linear least-squares fit

Linear least-squares fit is a problem of fitting n data points $\{x_i, y_i \pm \sigma_i\}$, where σ_i are experimental errors, by a linear combination of m functions

$$F(x) = \sum_{k=1}^{m} c_k f_k(x) \;. \tag{5}$$

The least-squares fit has to minimize the square deviation, called χ^2,

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{y_i - F(x_i)}{\sigma_i}\right)^2 \;. \tag{6}$$

Minimization of χ^2 with respect to the coefficiendt c_k in (5) is apparently equivalent to the least-squares problem (2) where

$$A_{ik} = \frac{f_k(x_i)}{\sigma_i} \;,\; b_i = \frac{y_i}{\sigma_i} \;. \tag{7}$$

The formal solution is

$$\mathbf{c} = R^{-1}Q^T\mathbf{b} \;, \tag{8}$$

however in practice it is better to back-substitute the system $R\mathbf{c} = Q^T\mathbf{b}$.

1.1.1 Errors and correlations

Suppose δy_i is a (small) deviation of the measured value of the physical variable from its exact value. The corresponding deviation δc_k of the fitting coefficient is then given as

$$\delta c_k = \sum_i \frac{\partial c_k}{\partial y_i} \delta y_i \;. \tag{9}$$

In a good experiment the deviations δy_i are statistically independent and distributed normally with the standard deviations σ_i. The deviations (9) are then also distributed normally with *variances*,

$$\langle \delta c_k \delta c_k \rangle = \sum_i \left(\frac{\partial c_k}{\partial y_i}\sigma_i\right)^2 = \sum_i \left(\frac{\partial c_k}{\partial b_i}\right)^2 \;. \tag{10}$$

The standard errors in the fitting coefficients are then given as

$$\delta c_k = \sqrt{\langle \delta c_k \delta c_k \rangle} \;. \tag{11}$$

The variances are diagonal elements of the *covariance matrix*, Σ, made of *covariances*,

$$\Sigma_{kq} \equiv \langle \delta c_k \delta c_q \rangle = \sum_i \frac{\partial c_k}{\partial b_i}\frac{\partial c_q}{\partial b_i} \;. \tag{12}$$

Covariances $\langle \delta c_k \delta c_q \rangle$ are measures of how the coefficients c_k and c_q change together if the measured values y_i are varied. The normalized covariances,

$$\langle \delta c_k \delta c_q \rangle \big(\langle \delta c_k \delta c_k \rangle \langle \delta c_q \delta c_q \rangle \big)^{-1/2}$$

are called *correlations.*

Using (12) and (8) the covariance matrix can be calculated as

$$\begin{aligned} \Sigma &= \left(\frac{\partial \mathbf{c}}{\partial \mathbf{b}}\right)\left(\frac{\partial \mathbf{c}}{\partial \mathbf{b}}\right)^T \\ &= R^{-1}(R^{-1})^T = (R^TR)^{-1} = (A^TA)^{-1} \;. \end{aligned} \tag{13}$$

In conclusion, the square roots of the diagonal elements of the covariance matrix (12) give the estimates of the errors of the fitting coefficients; the normalized off-diagonal elements give their correlations.

1.2 JavaScript implementation

```
function lsfit(x,y,dy,funs){
// Linear least squares fit
// uses: qrdec, qrback, inverse
// input: data points {x,y,delta_y};
    functions funs
// output: fitting coefficients c and
    covariance matrix S
  var dot = function(a,b)
    {var s=0;for(i in a)s+=a[i]*b[i];return
         s}
  var t_times = function(A,B)
    [[dot(A[r],B[c]) for(r in A)] for(c in
        B)];
  var A=[[funs[k](x[i])/dy[i] for(i in x)]
      for(k in funs)];
  var b=[y[i]/dy[i] for(i in y)];
  var [Q,R]=qrdec(A);
  var c=qrback(Q,R,b);
  var S=inverse(t_times(R,R))

  return [c,S]
}
```