Solving Least Squares Problems via the QR factorization[−1−].

The Problem. Suppose we are given *data*

$$(t_i, b_i), \quad i = 1, \cdots, m \tag{1}$$

and *basis functions*

$$\phi_j, \quad j = 1, \cdots, n \tag{2}$$

and we wish to find *coefficients*

$$x_j, \quad j = 1, \cdots, n \tag{3}$$

such that the approximating function

$$p(t) = \sum_{j=1}^{n} x_i \phi_j(t) \tag{4}$$

minimizes the sum of squares:

$$F(x_1, x_2, \cdots, x_n) = \sum_{k=1}^{m} \big(b_k - p(b_k)\big)^2 = \sum_{k=1}^{m} \left(b_k - \sum_{j=1}^{n} x_j \phi_j(t_k)\right)^2 = \min. \tag{5}$$

Purpose. This problem arises frequently in applications. The purpose of this note is to explain the most powerful and widely used technique for solving (5). It ends with a pointer to software that presents the most sophisticated implementation of the mathematical ideas described here.

Matrix Formulation. First observe that if we define

$$A = [\phi_j(t_i)]_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}} \in \mathbb{R}^{m \times n}, \quad x = [x_j]_{j=1,\cdots,n} \in \mathbb{R}^n \quad \text{and} \quad b = [b_i]_{i=1,\cdots,m} \in \mathbb{R}^m \tag{6}$$

then (5) is equivalent to

$$\boxed{\text{Find } x \in \mathbb{R}^n \quad \text{such that} \quad F(x) = \|Ax - b\|_2 = \min.} \tag{7}$$

We assume that

$$m \geq n \quad \text{and} \quad \text{rank} A = n. \tag{8}$$

The Normal Equations. Differentiating with respect to the x_i in (7) or (5) shows that the solution x satisfies the normal equations

$$\boxed{A^T A x = A^T b.} \tag{9}$$

The coefficient matrix $A^T A$ is symmetric and positive definite. However, solving (9) directly is poorly behaved with respect to round-off effects since the condition number of $A^T A$ is the square of that of A. In this note we consider a superior approach. The word *normal* refers to the fact that the difference $b - Ax$ in (9) is perpendicular (normal) to the column space of A.

[−1−] by Peter Alfeld, 801-581-6842, JWB 127, pa@math.utah.edu. TEX processed on September 25, 2013.

Problem (5) is called a *discrete linear least squares problem.* “Discrete” because there are finitely many data, “least because we minimize “squares”, and “linear” because we solve a linear system.

Example. Suppose

$$n = 2, \quad \phi_1(t) = t, \quad \text{and} \quad \phi_2(t) = 1. \tag{10}$$

Thus we approximate our data by a linear function

$$p(t) = x_1 t + x_2 \tag{11}$$

and whose graph is a straight line. This special case is also known as *linear regression.* The normal equations become

$$\begin{pmatrix} \sum_{k=1}^{m} t_k^2 & \sum_{k=1}^{m} t_k \\ \sum_{k=1}^{m} t_k & \sum_{k=1}^{m} 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} b_k t_k \\ \sum_{k=1}^{m} b_k \end{pmatrix} \tag{12}$$

In this example, the world “linear” refers to the approximating linear function.

The QR approach. Suppose we write

$$\boxed{A = QR} \tag{13}$$

where

$$Q = \begin{array}{cc} & \begin{array}{cc} n & m-n \end{array} \\ m & \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \end{array} \tag{14}$$

is *orthogonal* and

$$R = \begin{array}{cc} & n \\ \begin{array}{c} n \\ m-n \end{array} & \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \end{array} \tag{15}$$

with R_1 being upper triangular. An orthogonal matrix Q is one that satisfies

$$Q^{-1} = Q^T. \tag{16}$$

Obviously, if Q is orthogonal, so is Q^T. An orthogonal matrix is never singular (why), its condition number (w.r.t. $\|\cdot\|_2$) is 1, and it is not always symmetric. A significant property of an orthogonal matrix is that multiplying with it does not alter the (Euclidean) norm of a vector:

$$\|Qx\|_2^2 = (Qx)^T(Qx) = x^T Q^T Q x = x^T x = \|x\|_2^2. \tag{17}$$

Thus the first n columns of Q form an orthonormal basis of the column space of A. We obtain

$$\begin{aligned} \|Ax - b\|_2^2 &= \|Q^T(Ax - b)\|_2^2 \\ &= \|Q^T Ax - Q^T b\|_2^2 \\ &= \left\| \begin{pmatrix} R_1 x \\ 0 \end{pmatrix} - \begin{pmatrix} Q_1^T b \\ Q_2^T b \end{pmatrix} \right\|_2^2 \\ &= \left\| R_1 x - Q_1^T b \right\|_2^2 + \left\| Q_2^T b \right\|_2^2. \end{aligned} \tag{18}$$

Of the two terms on the right we have no control over the second, and we can render the first one zero by solving (the square triangular $n \times n$ linear system)

$$R_1 x = Q_1^T b. \tag{19}$$

Note that we do not use A^TA with it's squared condition number, and of course we don't have to calculate Q_2.

Computation of the QR factorization. Actual Computation of the QR factorization is based on

Householder Reflections. Given a vector

$$u \in \mathbb{R}^m : \quad \|u\|_2 = 1 \tag{20}$$

the corresponding Householder reflection (or Householder matrix) H is defined by

$$\boxed{H = I - 2uu^T \quad \text{with} \quad u^Tu = 1.} \tag{21}$$

The following properties of H are easily verified:

1. It is symmetric, i.e., $H^T = H$.
2. It is orthogonal, i.e., $H^{-1} = H^T$.
3. It projects vectors through the (hyper-)plane orthogonal to u. Specifically:

$$Hu = u - 2uu^Tu = u - 2u = -u \tag{22}$$

and

$$u^Tv = 0 \quad \Longrightarrow \quad Hv = v - 2uu^Tv = v. \tag{23}$$

Thus u gets transformed into its negative and a vector v in the plane is preserved.

Zeroing a column. The key ingredients in computing the QR factorization are Householder reflections H that take a vector a (which basically will be a column of A) and take it to a vector Ha that is zero below the first entry. Let e denote the vector that is 1 in the first entry, and zero everywhere else. We write u as

$$\boxed{u = \frac{v}{\|v\|_2}.} \tag{24}$$

Thus we first find a vector v that has the right direction, and then we normalize it. Since multiplication with an orthogonal matrix does not alter the norm of a vector we have to have

$$Ha = (I - 2uu^T)a = a - (2u^Ta)\frac{v}{\|v\|_2} = \pm\|a\|_2 e. \tag{25}$$

So clearly, v must be a linear combination of a and e. So we let

$$v = a + \alpha e \tag{26}$$

where α is as yet unknown. **(Thus we have reduced the problem from finding an unknown vector to the problem of finding an unknown scalar.)** Letting a_1 denote the first entry of A, observe that

$$v^Tv = a^Ta + 2\alpha a_1 + \alpha^2 \tag{27}$$

and

$$Ha = a - \frac{2(a+\alpha e)^Ta}{a^Ta + 2\alpha a_1 + \alpha^2}(a + \alpha e) = \left(1 - \frac{2(a+\alpha e)^Ta}{a^Ta + 2\alpha a_1 + \alpha^2}\right)a - \frac{2\alpha v^Ta}{v^Tv}e. \tag{28}$$

Since Ha is a multiple of e the coefficient of a must vanish, i.e.,

$$2a^Ta + 2\alpha a_1 = a^Ta + 2\alpha a_1 + \alpha^2. \tag{29}$$

This gives

$$\|a\|_2^2 = \alpha^2, \tag{30}$$

i.e.,

$$\boxed{\alpha = \pm\|a\|_2}. \tag{31}$$

Thus

$$u = \frac{a \pm \|a\|_2 e}{\big\|a \pm \|a\|_2 e\big\|_2}. \tag{32}$$

Since the purpose of doing these calculations in such detail is to illustrate common techniques it is worthwhile to check directly that this choice works.

We obtain

$$\big\|a \pm \|a\|_2 e\big\|_2^2 = (a \pm \|a\|_2 e)^T (a \pm \|a\|_2 e) = a^T a \pm 2\|a\|_2 a_1 + \|a\|_2^2 = 2\left(a^T a \pm \|a\|_2 a_1\right). \tag{33}$$

Thus

$$\begin{aligned}
Ha &= \left(I - 2\frac{(a \pm \|a\|_2 e)(a \pm \|a\|_2 e)^T}{\big\|a \pm \|a\|_2 e\big\|_2^2} \right) a \\
&= a - \frac{2(a \pm \|a\|_2 e)^T a}{\big\|a \pm \|a\|_2 e\big\|_2^2}(a \pm \|a\|_2 e) \\
&= a - \frac{2(a^T a \pm \|a\|_2 e^T a)}{\big\|a \pm \|a\|_2 e\big\|_2^2}(a \pm \|a\|_2 e) \\
&= a - \frac{2(a^T a \pm \|a\|_2 a_1)}{\big\|a \pm \|a\|_2 e\big\|_2^2}(a \pm \|a\|_2 e) \\
&= a - (a \pm \|a\|_2 e) \\
&= \mp\|a\|_2 e.
\end{aligned} \tag{34}$$

So how do we pick the sign of α? The usual choice is

$$\alpha = \operatorname{sign}(a_1)\,\|a\|_2.$$

This makes the first entry of u as large as possible (in absolute value) and tends to counteract the risk of dividing by a small number when normalizing v. So our final choice is

$$\boxed{u = \frac{a + \operatorname{sign}(a_1)\|a\|_2 e}{\big\|a + \operatorname{sign}(a_1)\|a\|_2 e\big\|_2}.} \tag{35}$$

It is also worth checking that the coefficient of e in (28) has absolute value equal to $\|a\|_2$, as it must since multiplication with an orthogonal matrix does not alter the 2-norm of a vector. Using (27) and (31) we have

$$\frac{2\alpha v^T a}{v^T v} = \frac{2\alpha(a^T a + \alpha a_1)}{a^T a + 2\alpha a_1 + \alpha^2} = \frac{2\alpha(a^T a + \alpha a_1)}{2a^T a + 2\alpha a_1} = \alpha = \pm\|a\|. \tag{36}$$

Putting it together. Let H_1 be the Householder reflection that zeros the first column of A below the diagonal. Thus

$$H_1 A = \begin{bmatrix} x & x \\ 0 & \hat{A} \end{bmatrix} \tag{37}$$

where $\hat{A}$ is an $m-1 \times n-1$ matrix (and x is a generic notation for non-zero entries). Let $\hat{H}_2$ be the $m-1 \times m-1$ Householder reflection that similarly reduces the first of column of $\hat{A}$ to zero below the diagonal. Then let

$$H_2 = \begin{bmatrix} 1 & 0^T \\ 0 & \hat{H}_2 \end{bmatrix} \tag{38}$$

Clearly

$$H_2 H_1 A = \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & A_3 \end{bmatrix}. \tag{39}$$

Continuing in this fashion we construct Householder reflections H_1, H_2, $\cdots$, H_n such that

$$H_n H_{n-1} \cdots H_2 H_1 A = R. \tag{40}$$

The matrix Q_1^T contains the first n rows of the product $H_n H_{n-1} \cdots H_2 H_1$. As a practical matter, of course we do not store the (full $m \times m$) matrices H_i, but rather only the vectors that define them.

Multiplying with Householder Reflections. Note that

$$(I - 2uu^T)A = A - 2u(u^T A). \tag{41}$$

Multiplication of A with a Householder reflection changes A by a matrix of rank 1. This can be implemented in $\mathcal{O}(n^2)$ operations. If the multiplication is implemented as an ordinary matrix multiplication $\mathcal{O}(n^3)$ operations are required, which would be grossly wasteful.

The Gram-Schmidt Process.

A better known method for computing the QR factorization is the Gram-Schmidt process. It is not as stable numerically as the Householder approach, as discussed in Golub and van Loan. However, it applies in more general inner product spaces and is worth knowing.

So consider again the factorization $A = QR$, where Q is orthogonal and R is triangular. It's clear that the first column of A is a multiple of the first column of Q. In general, the first k columns of A are linear combinations of the first k columns of Q, and vice versa. Thus the first k columns of Q span the same space as the first k columns of A. So we can think of computing Q as the first n steps in the following problem:

Given a sequence of linearly independent vectors

$$a_1, a_2, a_3, \ldots \tag{42}$$

construct an orthonormal sequence of vectors

$$q_1, q_2, q_3, \ldots \tag{43}$$

such that

$$q_i^T q_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases} \qquad \text{and} \qquad \operatorname{span}\{q_1, q_2, \ldots q_k\} = \operatorname{span}\{a_1, a_2, \ldots a_k\}, \quad k = 1, 2, 3, \ldots$$

This problem can be solved as follows:

1. Let

$$q_1 = \frac{a_1}{\|a_1\|} \tag{44}$$

2. For $k = 1, 2, \ldots$
   a. Define $v_k = a_k - \sum_{j=1}^{k-1} q_j^T a_k q_j$
   b. Let $q_k = \frac{v_k}{\|v_k\|}$.

It's clear by induction that this sequence has the desired properties since, for $i < k$,

$$v_k^T q_i = a_k^T q_i - \sum_{j=1}^{k-1} a_j^T a_k q_i^T q_j = a_k^T q_i - a_k^T q_i = 0. \tag{45}$$

References

1. E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson, LAPACK User's guide, 1992, SIAM publication, ISBN 0-89871-294-7.
2. B.S. Garbov, J.M. Boyle, J.J. Dongarra, C.B. Moler, Matrix Eigensystem Routines Eispack Guide Extension, Springer Lecture Notes in Computer Science, v. 51, Springer Verlag, 1977, ISBN 0-387-08254-9.
3. Gene H. Golub and Charles F. van Loan, Matrix Computations, 4th ed., The Johns Hopkins University Press, 2013, ISBN 13: 978-4214-0794-4.
4. Charles L. Lawson and Richard J. Hanson, Solving Least Squares Problems, Prentice Hall, 1974, ISBN 0-13-822585-0