Report on Numerical Linear Algebra Coding Assignment
Chapter 3: Least Squares Problem

Hongyi Zhang, 1500017736

November 8, 2017

1 QR factorization for least squares problem

1.1 Problem formulation

Two problems of the least squares kind are presented. For the first problem, we would like to find a quadratic polynomial $f(t) = at^2 + bt + c$ that best fits the given data $\{(t_i, y_i)\}_{i=1}^n$ in the sense of L_2 norm, which is namely minimizing the residual sum of squares

$$RSS = \sum_{i=1}^{n} [y_i - (at_i^2 + bt_i + c)]^2 = \|y - Ax\|_2^2,$$

where $y = (y_1, \cdots, y_n)^T$, the matrix

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ \cdots & \cdots & \cdots \\ 1 & t_n & t_n^2 \end{bmatrix},$$

and the vector $x = (c, b, a)^T$. Hence, it is reduced to a least squares problem.

Another problem requires a linear model for house price predictions, and formally a linear combination $y = x_0 + a_1x_1 + a_2x_2 + \cdots + a_{11}x_{11}$ of a series of variables that best fits the given data $\{(a_1^i, \cdots, a_{11}^i, y_i)\}_{i=1}^n$ in the sense of L_2 norm. Similar to the previous problem, the objective is to minimize $\|y - Ax\|_2$ where $y = (y_1, \cdots, y_n)^T$, the matrix

$$A = \begin{bmatrix} 1 & a_1^1 & a_2^1 & \cdots & a_{11}^1 \\ 1 & a_1^2 & a_2^2 & \cdots & a_{11}^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_1^n & a_2^n & \cdots & a_{11}^n \end{bmatrix},$$

and the vector $x = (x_0, x_1, \cdots, x_{11})^T$. Thus it can also be resolved by the least squares technique.

1.2 Householder method for QR factorization

QR factorization plays a crucial role in the least squares problem. It states that for every matrix $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, there exists an orthogonal matrix $Q \in \mathbb{R}^{m\times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m\times n}$ such that

$$A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1,$$

where $Q_1 \in \mathbb{R}^{m\times n}$ has orthogonal columns, and $R_1 \in \mathbb{R}^{n\times n}$ is an upper triangular square matrix.

QR factorization can be done with Householder transformations. Given a vector x, let $v = x - \|x\|_2 e_1$, then the Householder reflection defined by $H = 1 - 2ww^T$ where $w = v/\|v\|_2$ introduces zeros to all but the first entries of x, i.e. $Hx = \|x\|_2 e_1$. Applying this technique to each column of A in a way similar to Gaussian elimination, all entries below the main diagonal would vanish, yielding an upper triangular R and an orthogonal Q which is the composition of all the reflections.

Once the factorization is done, then the least squares can be rewritten as

$$\|y - Ax\|_2^2 = \|Q^T y - Q^T Ax\|_2 = \|Q_1^T y - R_1 x\|_2^2 + \|Q_2^T y\|_2^2.$$

Therefore, the solution to the upper triangular equation $R_1 x = Q_1^T y$ is exactly what we desire. In practice, Q_1 is not explicitly computed, and the product $Q_1^T y$ is obtained by successive application of the Householder reflections on the vector y.

1.3 Numerical results

For the first problem, the coefficient vector is $x^* = (1,1,1)^T$, and the corresponding best quadratic polynomial is

$$f^*(t) = t^2 + t + 1.$$

The norm of the residual vector $\|y - Ax^*\|$ is exactly zero, which is a fairly surprising result. But normally, we cannot expect a result as desirable as this one.

For the second problem, the optimal coefficients are shown in Table 1, of which the first entry x_0 is the intercept. The minimal residual norm is $\|y - Ax^*\| = 16.3404$. Computing the R-squared statistics gives us $R^2 = 0.9507$, implying that our linear model actually fits the data quite well: about 95 percent of the variance in y is predictable from the model.

Table 1: Coefficients for the least squares linear model

x_0	x_1	x_2	x_3	x_4	x_5
2.0775	0.7189	9.6802	0.1535	13.6796	1.9868
x_6	x_7	x_8	x_9	x_{10}	x_{11}
−0.9582	−0.4840	−0.0736	1.0187	1.4435	2.9028

2 QR factorization for solving linear systems

2.1 Objective

Now we are going to solve a couple of linear systems with QR factorization. For a given equation $Ax = b$, the solution (if exists) is always the minimizer of the residual norm $\|Ax - b\|_2$ since it cannot go below zero, and hence it would be sufficient to apply the QR-least-squares method to minimize $\|Ax - b\|_2$ to solve the equation, in spite of more computations needed for the procedure.

2.2 Numerical results

We try the QR method on three particular equations, and the performance comparisons among different solving routines are shown in Table 2 and Table 3. From these results, we may conclude that QR method enjoys greater numerical stability when faced with ill-conditioned problems. But in the meanwhile, the drawback is the additional computational cost.

A point worth mentioning is that, the result for case 1 turns out to be an NaN (Not a Number) for the initial implementation of the algorithm. It normalized the Householder vector to have a leading 1 by dividing its first entry, which introduces more error despite the reduction in storage requirement. After removing this normalization process, the solution comes back to normal.

Table 2: Performance comparisons (I)

Case 1	L^∞ error	Avg. runtime	Case 2	L^∞ error	Avg. runtime
LU	5.368e+08	0.00643 sec	Cholesky	2.664e−15	0.02688 sec
PLU	2.797e−06	0.00934 sec	LDL^T	3.553e−15	0.00448 sec
QR	4.441e−16	0.01753 sec	QR	7.105e−15	0.01838 sec

Table 3: Performance comparisons (II)

Case 3	L^∞ error	Avg. runtime
LU	195.959	0.00212 sec
PLU	376.105	0.00410 sec
Cholesky	43018000.193	0.01633 sec
LDL^T	67.534	0.00118 sec
Matlab	95.030	0.00051 sec
QR	74.850	0.00459 sec

*Case 1: Tridiagonal matrix of order 84

*Case 2: Tridiagonal symmetric positive definite matrix of order 100

*Case 3: Hilbert matrix of order 40