Jane Kim
Numerical Linear Algebra
Homework 4
19 February 2019

2. Solve the following three systems by the three different QR decomposition methods.

   See source code at my Github repository.

   (a) Let

   $$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 7 \\ 4 & 2 & 3 \\ 4 & 2 & 2 \end{bmatrix}.$$

   Solve the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (1, 1, ..., 1)^T$.

   The following solution was returned by all three QR decomposition methods:

   $$\mathbf{x} = \begin{bmatrix} 0.17904 \\ -0.38865 \\ 0.40611 \end{bmatrix}.$$

   The norm $\|A\mathbf{x} - \mathbf{b}\|$ was 0.52451 for all three methods.

   (b) Let

   $$A = \begin{bmatrix} 0.70000 & 0.70711 \\ 0.70001 & 0.70711 \end{bmatrix}.$$

   Solve the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (1, 1)^T$.
   Both Gram-Schmidt methods returned

   $$\mathbf{x} = \begin{bmatrix} -4.6029\text{e-}6 \\ 1.41421 \end{bmatrix},$$

   with $\|A\mathbf{x} - \mathbf{b}\| = 3.2547\text{e-}11$, while the Householder method returned

   $$\mathbf{x} = \begin{bmatrix} -9.8648\text{e-}12 \\ 1.414207 \end{bmatrix},$$

   with $\|A\mathbf{x} - \mathbf{b}\| = 1.11022\text{e-}16$.

(c) Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 9 \end{bmatrix}.$$

Solve the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (6, 15)^T$.

The full QR decomposition was performed on A^* instead of A using the Householder transform based method. The resulting solution was

$$\mathbf{x} = \begin{bmatrix} 0.42857 \\ 0.85714 \\ 1.2857 \end{bmatrix},$$

with $\|A\mathbf{x} - \mathbf{b}\| = 3.66205\text{e-}15$.

3. Consider the following over-determined linear system:

$$A\mathbf{x} = \mathbf{b},$$

where $A \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. The Tikhonov regularization can be formulated as a minimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\|_2^2 + \alpha^2\|\mathbf{x}\|_2^2, \tag{1}$$

where $\alpha \geq 0$ is a non-negative number.

(a) Prove that the necessary condition to minimize the above function is $(A^TA + \alpha^2 I)\mathbf{x} = A^T\mathbf{b}$.

Proof. Let $F(\mathbf{x})$ be defined as the function we want to minimize:

$$F(\mathbf{x}) = \|\mathbf{b} - A\mathbf{x}\|_2^2 + \alpha^2\|\mathbf{x}\|_2^2 = \sum_{i=1}^{m}(b_i - (A\mathbf{x})_i)^2 + \alpha^2\sum_{j=1}^{n} x_j^2.$$

Taking the derivative of $F(\mathbf{x})$ with respect to the kth component of $\mathbf{x}$ gives

$$\begin{aligned}
\frac{\partial F(\mathbf{x})}{\partial x_k} &= \frac{\partial}{\partial x_k}\left\{\sum_{i=1}^{m}(b_i - (A\mathbf{x})_i)^2 + \alpha^2\sum_{j=1}^{n} x_j^2\right\} \\
&= -2\sum_{i=1}^{m}(b_i - (A\mathbf{x})_i)\left(\frac{\partial}{\partial x_k}(A\mathbf{x})_i\right) + 2\alpha^2 x_k
\end{aligned}$$

Since we can write $(A\mathbf{x})_i = \sum_{j=1}^{n} A_{ij}x_j$, we have

$$\begin{aligned}
\frac{\partial F(\mathbf{x})}{\partial x_k} &= -2\sum_{i=1}^{m}(b_i - (A\mathbf{x})_i)\left(\frac{\partial}{\partial x_k}\sum_{j=1}^{n} A_{ij}x_j\right) + 2\alpha^2 x_k \\
&= -2\sum_{i=1}^{m}(b_i - (A\mathbf{x})_i)A_{ik} + 2\alpha^2 x_k \\
&= 2\left(\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ik}A_{ij}x_j - \sum_{i=1}^{m} A_{ik}b_i + \alpha^2 x_k\right)
\end{aligned}$$

Then $F(\mathbf{x})$ is minimized with respect to x_k when

$$\begin{aligned}
0 &= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ik}A_{ij}x_j - \sum_{i=1}^{m} A_{ik}b_i + \alpha^2 x_k \\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{m}(A^T)_{ki}A_{ij}\right)x_j - \sum_{i=1}^{m}(A^T)_{ki}b_i + \alpha^2 x_k \\
&= \sum_{j=1}^{n}(A^TA)_{kj}x_j - (A^T\mathbf{b})_k + \alpha^2 x_k \\
&= (A^TA\mathbf{x})_k - (A^T\mathbf{b})_k + \alpha^2 x_k.
\end{aligned}$$

Therefore, the necessary condition to minimize $F(\mathbf{x})$ is

$$(A^TA + \alpha^2 I)\mathbf{x} = A^T\mathbf{b} \tag{2}$$

(b) Show that the above system can be rewritten as $\begin{bmatrix} A \\ \alpha I \end{bmatrix}\mathbf{x} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}$.

Proof. Define $B = \begin{bmatrix} A \\ \alpha I \end{bmatrix}$. Notice that the left-hand side of equation (2) is given by

$$B^TB\mathbf{x} = \begin{bmatrix} A \\ \alpha I \end{bmatrix}^T \begin{bmatrix} A \\ \alpha I \end{bmatrix}\mathbf{x} = \begin{bmatrix} A^T & \alpha I \end{bmatrix}\begin{bmatrix} A \\ \alpha I \end{bmatrix}\mathbf{x} = (A^TA + \alpha^2 I)\mathbf{x},$$

and the right-hand side is given by

$$B^T\begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} = \begin{bmatrix} A \\ \alpha I \end{bmatrix}^T\begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} = \begin{bmatrix} A^T & \alpha I \end{bmatrix}\begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} = A^T\mathbf{b}.$$

Thus we have that $B^TB\mathbf{x} = B^T\begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}$, or in other words,

$$B\mathbf{x} = \begin{bmatrix} A \\ \alpha I \end{bmatrix}\mathbf{x} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}. \tag{3}$$

(c) Apply the QR method with the above regularization to solve the Hilbert system, $A\mathbf{x} = \mathbf{b}$, where A is the nth order Hilbert matrix and $\mathbf{b} = (1, 1, ..., 1)^T$.

See source code at my Github repository.

(d) Observe the performance of the above method for $n \in \{10, 20, 30\}$ and $\alpha \in \{0.00000001, 0.00001, 0.001, 0.1\}$.

See Table 1.

(e) Can you choose an α which you think is ideal so that you get a best solution in the sense that is it closest to the true solution in the 2-norm?

Table 1: Comparison of solutions to the Hilbert system, using three different QR methods with Tikhonov regularization.

		classical Gram-Schmidt		modified Gram-Schmidt		Householder	
n	α	$\|A\mathbf{x}-\mathbf{b}\|$	$\|\mathbf{x}\|$	$\|A\mathbf{x}-\mathbf{b}\|$	$\|\mathbf{x}\|$	$\|A\mathbf{x}-\mathbf{b}\|$	$\|\mathbf{x}\|$
10	1e-8	0.004655	195872.	0.000187	10278.3	0.000187	10278.1
	1e-5	0.007126	394.067	0.007126	394.062	0.007126	394.062
	1e-3	0.068282	54.4000	0.068282	54.4000	0.068282	54.4000
	1e-1	0.714407	4.86861	0.714407	4.86861	0.714407	4.86861
20	1e-8	0.170031	13358994	0.000332	25971.7	0.000332	25971.0
	1e-5	0.009959	613.559	0.0099598	613.559	0.0099598	613.559
	1e-3	0.104120	71.7779	0.104120	71.7779	0.104120	71.7779
	1e-1	1.04718	7.18780	1.04718	7.18780	1.04718	7.18780
30	1e-8	1.82094	65965561	0.00040641	27787.5	0.00040641	27786.9
	1e-5	0.012149	932.273	0.012145	931.649	0.012145	931.649
	1e-3	0.122509	91.9022	0.122509	91.9022	0.122509	91.9022
	1e-1	1.27993	9.15816	1.27993	9.15816	1.27993	9.15816

In an ideal world, the solution as $\alpha \to 0$ is closest to the true solution in the 2-norm. However, the data in Table 1 shows that too small of an α causes $\mathbf{x}$ to grow to a potentially unmanageable size. In addition, there is also a significant difference in stability between the classical and modified Gram-Schmidt methods for $\alpha =$ 1e-8, suggesting that a small α can lead to a large accumulation of rounding error. The solutions are most stable, both across order n and across methods, for $\alpha \geq$ 1e-5. Therefore, I think $\alpha =$1e-5 would be the best choice.