Midterm Exam

Shuowen Wei

October 27, 2011

Problem 1

Proof:

By the given conditions, we can easily get: $Ir + Ax = b$ and $A^T r = 0$. Then left multiply A^T to the first equation, we have,

$$A^T Ir + A^T Ax = 0 + A^T Ax = A^T b$$

thus $A^T Ax = A^T b$, since A is full rank, then by **Theorem 11.2**, we know that the solution x of $A^T Ax = A^T b$ minimize the $||Ax - b||_2$.

Problem 2

Solution:

Since A is a real matrix with reduced SVD $A = \hat{U}\hat{\Sigma}V^T$, so $\hat{U}$ is m-by-n real unitary matrix and $\Sigma,\hat{} V$ are both n-by-n real matrix with is Σ diagonal and V is unitary.

Thus, we have:

1.

$$(A^T A)^{-1} A^T = (V\hat{\Sigma}\hat{U}^T\hat{U}\hat{\Sigma}V^T)^{-1}(\hat{U}\hat{\Sigma}V^T)^T = (V\hat{\Sigma}^{-2}V^T)(V\hat{\Sigma}\hat{U}^T) = V\hat{\Sigma}^{-1}\hat{U}^T(= A^{-1})$$

2.

$$A(A^T A)^{-1} = (\hat{U}\hat{\Sigma}V^T)(V\hat{\Sigma}\hat{U}^T\hat{U}\hat{\Sigma}V^T)^{-1} = (\hat{U}\hat{\Sigma}V^T)(V\hat{\Sigma}^{-2}V^T) = \hat{U}\hat{\Sigma}^{-1}V^T(= (A^T)^{-1})$$

Problem 3

Proof:

For an arbitrary real matrix $A \in \mathbb{C}^{m\times n}$, Moore-Penrose pseudoinverse of A (A must be full rank), denoted by A^+, is defined as:

$$A^+ = (A^T A)^{-1} A^T \in\in \mathbb{C}^{n\times m}$$

Since real matrix A has a reduced SVD $A = \hat{U}\hat{\Sigma}V^T$, then by using the results of problem 2, we verify:

$$AA^+A = A((A^T A)^{-1}A^T)A = (\hat{U}\hat{\Sigma}V^T)(V\hat{\Sigma}^{-1}\hat{U}^T)(\hat{U}\hat{\Sigma}V^T) = \hat{U}\hat{\Sigma}V^T = A$$

and

$$A^+A = ((A^TA)^{-1}A^T)A = (V\hat{\Sigma}^{-1}\hat{U}^T)(\hat{U}\hat{\Sigma}V^T) = I = I^T = (A^+A)^T$$

Thus the Moore-Penrose pseudoinverse of A satisfies those two identities.

Problem 4

Proof:

It is very similar to **Prop 11.2**, let $Az = 0$, then if x minimizes $||Ax - b||_2$, so does $x + z$. And z is any element of $null(A)$, since $A \in \mathbb{C}^{m\times n}$ is full rank, then $rank(A) = m$ since $m < n$. So $dim(null(A)) = n - m$, thus the solution $x + z$ is an (n-m)-dimentional set.

Since $A \in \mathbb{C}^{m\times n}(m < n)$, so $A^*A \in \mathbb{C}^{n\times n}$ is not nonsingular any more because its rank is m and $m < n$. But $AA^* \in \mathbb{C}^{m\times m}$ is nonsingular and invertible. So it is nature to think of another way to set $x = A^*y$ such that $AA^*y = b$. Once we compute y, we get x. And it is easy to find that $y = (AA^*)^{-1}b$, thus one solution of the underdetermined least square problem is:

$$x = A^*y = A^*(AA^*)^{-1}b$$

First of all, we will prove that this solution is exactly the unique minimum norm solution.

Suppose that for an arbitrary solution X of this problem, i.e., $AX = b$, then we have $A(X - x) = 0$. Thus

$$(X - x)^*x = (X - x)^*A^*(AA^*)^{-1}b = (A(X - x))^*(AA^*)^{-1}b = 0$$

then $X-x$ and x are mutually orthogonal, so we have $||X||^2 = ||X-x+x||^2 = ||X-x||^2 + ||x||^2 \geqslant ||x||^2$, since X is the arbitrary solution of this problem, then x is the unique minimum norm solution.

Secondly, we will give out different algorithms to compute this mim-norm solution:

Appropriately modified normal equations:

1. Form AA^*
2. Compute the Cholesky factorization $AA^* = R^*R$
3. Solve the lower-triangular system $R^*w = b$ for w
4. Solve the upper-triangular system $Rz = w$ for z
5. Set $x = A^*z$

QR decomposition:

1. Compute the reduced QR factorization $A^* = \hat{Q}\hat{R}$, where $\hat{Q} \in \mathbb{C}^{n\times m}, \, and \, \hat{R} \in \mathbb{C}^{m\times m}$, $\hat{Q}$ is unitary and $\hat{R}$ is nonsingular, upper triangular
2. Solve the lower-triangular system $\hat{R}^*z = b$ for z
3. Set $x = \hat{Q}z$

SVD:

1. Compute the reduce SVD $A^* = \hat{U}\hat{\Sigma}V^*$
2. Compute the vector $\hat{V}^*b$
3. Solve and $\hat{\Sigma}w = \hat{V}^*b$ for w
4. Set $x = \hat{U}w$

Problem 11.1

Proof:

Since $A_1$is nonsingular matrix of dimension $n \times n$, then A is full rank, so by the definition of pseudoinverse of A, denoted by A^+, $A^+ = (A^*A)^{-1}A^*$. And A has reduced SVD $A = \hat{U}\hat{\Sigma}V^*$, then $A^*A = V\Sigma^2V^*$, so we have

$$||A^+||_2 = ||(A^*A)^{-1}A^*||_2 = ||V\Sigma^{-2}V^*V\Sigma U^*||_2 = ||V\Sigma^{-2}V^*V\Sigma U^*||_2 = ||\Sigma^{-1}||_2 = \frac{1}{\sigma_n}$$

where σ_n is the smallest singular value, because σ_1, σ_2, ...σ_n are in decreasing order.

Since $V\Sigma^2V^* = A^*A = \begin{bmatrix} A_1^* & A_2^* \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = A_1^*A_1 + A_2^*A_2$, then

$$\Sigma^2 = V^*(A_1^*A_1 + A_2^*A_2)V = V^*A_1^*A_1V + V^*A_2^*A_2V = \Sigma_1^2 + \Sigma_2^2$$

where Σ_1 and Σ_2 are the singular value matrix from the SVD of $A_1$and A_2.

Because A_1 is full rank, thus every element in Σ_1^2 is positive, denote by α_i^2, $i = 1, 2...n$. And denote the element in Σ_2^2 as β_i^2, $i = 1, 2...n$. Both α_i^2 and β_i^2 are in decreasing order. So

$$\sigma_i^2 = \alpha_i^2 + \beta_i^2,\ i = 1, 2...n$$

Since A_2 is arbitrary, then $\beta_i^2 \geq 0$, (i.e some β_i^2 may equals to zero), thus $\sigma_i^2 \geq \alpha_i^2$, $i = 1, 2...n$, let $i = n$, then $\frac{1}{\sigma_n} \leq \frac{1}{\alpha_n} = ||A_1^{-1}||_2$, hence

$$||A^+||_2 \leq ||A_1^{-1}||_2$$

Problem 11.3

Please see the m-file: **midterm.m**

Just copy the whole code and run it, we will get as follows:

the solution 1 is very good and acceptable
the solution 2 is very good and acceptable
the solution 3 is very good and acceptable
the solution 4 is very good and acceptable
the solution 5 is very good and acceptable
the solution 6 is very good and acceptable

===== Please input "result" directly in the Command Window to get all the solutions if you like =====

Below are the six lists of the twelve coefficients:

coeff = Columns 1 through 3

999.999961392580e-003 1.00000000166766e+000 1.00000000099660e+000
934.609598937149e-003 934.609643543441e-003 934.609642735242e-003
746.990399160323e-003 746.990362478313e-003 746.990363182523e-003
461.679130551934e-003 461.679153132704e-003 461.679152633007e-003
115.989201689278e-003 115.989207162564e-003 115.989207125290e-003

-244.869859492343e-003 -244.869886874760e-003 -244.869886336078e-003
-573.704742484316e-003 -573.704721069467e-003 -573.704721555216e-003
-827.510048967594e-003 -827.510044386540e-003 -827.510044432986e-003
-973.092981738117e-003 -973.093006396147e-003 -973.093005888365e-003
-991.414199474889e-003 -991.414170358503e-003 -991.414170994523e-003
-880.077444206421e-003 -880.077476583473e-003 -880.077475870130e-003
-653.643593875715e-003 -653.643620146207e-003 -653.643619561760e-003
Columns 4 through 6
1.00000000099661e+000 1.00000000099661e+000 1.00000000099661e+000
934.609642735245e-003 934.609642735245e-003 934.609642735244e-003
746.990363182526e-003 746.990363182526e-003 746.990363182525e-003
461.679152633008e-003 461.679152633008e-003 461.679152633008e-003
115.989207125290e-003 115.989207125291e-003 115.989207125290e-003
-244.869886336077e-003 -244.869886336077e-003 -244.869886336078e-003
-573.704721555216e-003 -573.704721555216e-003 -573.704721555216e-003
-827.510044432986e-003 -827.510044432985e-003 -827.510044432987e-003
-973.093005888366e-003 -973.093005888364e-003 -973.093005888366e-003
-991.414170994525e-003 -991.414170994522e-003 -991.414170994525e-003
-880.077475870132e-003 -880.077475870130e-003 -880.077475870132e-003
-653.643619561763e-003 -653.643619561760e-003 -653.643619561764e-003

The mutual differences between the observations are very small, such that we can even ignor the differences and think they are the same. And the normal equations, precisely speaking, exihits a little instability, but its solution can be acceptable.