Mathematics 372 – Numerical Linear Algebra
Selected Solutions – Problem Set 6
March 7, 2007

3.5.24. a) If $x \in \mathbf{R}^n$ is any vector then

$$x^t(A^tA)x = (Ax)^t(Ax) = \|Ax\|_2^2.$$

This is always ≥ 0 by the properties of vector norms. Hence A^tA is positive *semi-definite* for all A.

b) If $A \in M_{n\times m}(\mathbf{R})$ and $n < m$, then there are fewer equations than variables in the homogeneous system $Ax = 0$. This means that there is a nontrivial solution $x_p \neq 0$ with $Ax_p = 0$ (set at least one of the free variables to a nonzero value). Hence

$$x_p(A^tA)x_p = \|Ax_p\|_2^2 = 0$$

but $x_p \neq 0$. This shows that A^tA is *not positive-definite.*

c) Now assume $A \in M_{n\times m}(\mathbf{R})$ and $n \geq m$. We want to show that A^tA is positive definite if and only if A has full rank. (That is, $\text{rank}(A) = m$, or equivalently, that the columns of A are linearly independent – recall that the rank of a matrix is, by definition, the dimension of its column space.) By the same reasoning as above, we have A^tA is positive definite if and only if $x^t(A^tA)x = \|Ax\|_2^2 = 0$ implies $x = 0$, so $Ax = 0$ implies $x = 0$. But if $x \in \mathbf{R}^m$, then

$$Ax = x_1a_1 + \cdots + x_ma_m,$$

where a_i are the columns of A. Hence $Ax = 0$ implies $x = 0$ is equivalent to the assertion that the columns of A are linearly independent.

3.5.26. Note: A, B here are $n \times m$ matrices. a) $\Rightarrow$: Assume that x is the least-squares solution of $Ax = b$. Then we know $b - Ax \in R(A)^\perp = N(A^t)$ by Corollary 3.5.20 in the text. But we are assuming $R(B) = R(A)$ so by Theorem 3.5.10:

$$N(B^t) = R(B)^\perp = R(A)^\perp = N(A^t).$$

Therefore $B^t(b - Ax) = 0$. Rearranging algebraically, we get $B^tAx = B^tb$.

$\Leftarrow$: Conversely, suppose $B^tAx = B^tb$. Then $B^t(b - Ax) = 0$, so $b - Ax \in N(B^t) = R(B)^\perp$. Since we assume $R(A) = R(B)$, $R(A)^\perp = R(B)^\perp$, so $b - Ax \in R(A)^\perp$. Hence x is the least-squares solution of $Ax = b$ by Corollary 3.5.20.

b) $\Leftarrow$: Assume $A = BC$ for some nonsingular $m \times m$ matrix C. This matrix equation shows that each column of A is a linear combination of the columns of B so $R(A) \subseteq R(B)$ (using 3.5.13 part a). But C is nonsingular, so we also know $B = AC^{-1}$. Then the same reasoning shows that every column of B is a linear combination of the columns of A, so $R(B) \subseteq R(A)$. The two inclusions show $R(A) = R(B)$.

⇒: We must show that if $R(A) = R(B)$, then there is a nonsingular matrix C such that $A = BC$. First, note that $R(A) = R(B)$ implies $\text{rank}(A) = \text{rank}(B)$. Let us consider the special case first where the first k columns of A form a basis for $R(A)$ and similarly the first k columns of B form a basis for $R(B)$. Write $A = (\hat{A}\ A')$ and $B = (\hat{B}\ B')$ where $\hat{A}$ and $\hat{B}$ are the first k columns in each case. Since the columns of $\hat{A}$ and $\hat{B}$ are bases of $R(A) = R(B)$, each of the first k columns of A can be written uniquely as a linear combination of the first k columns of B, and vice versa. This says that there are $k \times k$ matrices $\hat{C}$ and $\hat{D}$ such that

$$\hat{A}\hat{D} = \hat{B} \tag{1}$$

and

$$\hat{B}\hat{C} = \hat{A}. \tag{2}$$

It follows that $\hat{A}(\hat{D}\hat{C}) = \hat{A}$ and $\hat{B}(\hat{C}\hat{D}) = \hat{B}$. Then, since the columns of $\hat{A}$ and $\hat{B}$ are linearly independent, we get $\hat{C}\hat{D} = I_k$ and $\hat{D}\hat{C} = I_k$, so $\hat{C}$ and $\hat{D}$ are inverse matrices of each other.

Furthermore, since the first k columns of B form a basis for $R(B)$ again, the last $m-k$ columns of A (the matrix A') can be written in terms of the entries in the matrix B in the form

$$A' = B' + \hat{B}P \tag{3}$$

for some $k \times (m-k)$ matrix P. Similarly,

$$B' = A' + \hat{A}Q \tag{4}$$

for some $k \times (m-k)$ matrix Q. But then consider the $m \times m$ matrices

$$C = \begin{pmatrix} \hat{C} & P \\ 0 & I_{m-k} \end{pmatrix}$$

and

$$D = \begin{pmatrix} \hat{D} & Q \\ 0 & I_{m-k} \end{pmatrix}$$

We have

$$BC = (\,\hat{B} \quad B'\,) \begin{pmatrix} \hat{C} & P \\ 0 & I_{m-k} \end{pmatrix} = (\,\hat{A} \quad A'\,) = A$$

by (2) and (3) above. Furthermore, C is invertible since $\hat{C}$ is invertible as above.

(In fact, although this is not necessary for the problem, it can be seen that $C^{-1} = D$. This is true since

$$B(CD) = (\,\hat{B} \quad B'\,) \begin{pmatrix} \hat{C} & P \\ 0 & I_{m-k} \end{pmatrix} \begin{pmatrix} \hat{D} & Q \\ 0 & I_{m-k} \end{pmatrix} = (\,\hat{B}(\hat{C}\hat{D}) \quad \hat{B}\hat{C}Q + \hat{B}P + B'\,).$$

From (2), (3), and (4),

$$\hat{B}\hat{C}Q + \hat{B}P + B' = \hat{A}Q + \hat{B}P + B' = B' - A' + A' - B' + B' = B'$$

But this shows

$$\hat{B}(\hat{C}Q + P) = 0,$$

so, since the columns of $\hat{B}$ are linearly independent, $\hat{C}Q + P = 0$. Hence $CD = I_m$, and similarly $DC = I_m$.)

The general case now follows since we can find some collections of columns of A and B that are bases for $R(A)$ and $R(B)$ respectively and reorder the columns to put those bases in the first k columns. This can be done by multiplying by suitable $m \times m$ permutation matrices Π, Σ. Let

$$A_1 = A\Pi \quad \text{and} \quad B_1 = B\Sigma.$$

Then by the above if $R(A) = R(A_1) = R(B_1) = R(B)$, there is an invertible $m \times m$ matrix C such that $B_1 C_1 = A_1$. But then

$$A = A_1 \Pi^{-1} = B_1 C_1 \Pi^{-1} = B(\Sigma C_1 \Pi^{-1}),$$

so $A = BC$ where $C = \Sigma C_1 \Pi^{-1}$. This is invertible since Π, Σ, C_1 are.