ACM 104

Homework Set 3 Solutions

January 29, 2001

2 Franklin Chapter 2, Problem 2, page 50.
Let A be an mxn matrix, and B be an $m \times p$ matrix. Give a necessary and sufficient condition so that the system $AX = B$ will have an $n \times p$ matrix-solution X. If X^0 is a particular solution, and if $\text{rank}(A) = r$, then describe all other solutions.

Similarly to theorem 2.6 on page 48, we have:
If A is an $m \times n$ matrix, and B is an mxp matrix, then $AX = B$ has a solution if the $m \times (n + p)$ matrix G formed by adjoining A to B has $\text{rank}(A) = \text{rank}(G)$. In this case the set of all solutions consists of any particular solution X^0 plus a linear space of dimension $(n - \text{rank}(A))p$.

Proof:

Let L_A be the linear space spanned by A and L_G be the linear space spanned by G. L_A can be spanned by $\text{rank}(A)$ linearly independent column vectors of A. Since $\text{rank}(A) = \text{rank}(G)$ the same column vectors also span L_G. That means that the column vectors in B can be written as a linear combination of the linearly independent column vectors of A and that proves sufficiency.
If $\text{rank}(A) = \text{rank}(G)$ does not hold then a column vector in B cannot be written as a linear combination of the column vectors of A, so the condition is necessary as well.

Let $\text{rank}(A) = \text{rank}(G) = r$. If X^0 is one solution of the equation $AX = B$ then for any other solution X we have:

$$A(X - X^0) = 0$$

By theorem 5 on page 41 the column vectors of $X - X^0$ are in an $n - r$ dimensional space, with basis say $v^1, \ldots, v^{n-r}$. If we look at matrices as an ordered set of column vectors, then a basis for the set of solution matrices X is:

$$\begin{gathered}(v^1, 0, \ldots, 0), (0, v^1, 0, \ldots, 0), \ldots, (0, \ldots, 0, v^1), \\ \cdots \\ (v^{n-r}, 0, \ldots, 0), (0, v^{n-r}, 0, \ldots, 0), \ldots, (0, \ldots, 0, v^{n-r})\end{gathered}$$

The dimension of the solution space is $(n - r)p$, and it is obviously a linear space.

3 Franklin Chapter 4, Problem 5, page 80.
If A is similar to B, show that the matrices A^k and B^k are similar for all powers $k = 0, 1, 2, \ldots$. If $det(A) \neq 0$, also show that $A^{-1}$is similar to B^{-1}.

A is similar to B means that there exists an invertible matrix T such that:

$$T^{-1}AT = B \tag{1}$$

Multiplying the left hand side of the equation by $T^{-1}AT$ and the right hand side with A we get:

$$\begin{array}{rcl} T^{-1}AT \ldots T^{-1}AT & = & B \ldots B \\ T^{-1}A^kT & = & B^k \end{array} \tag{2}$$

so A^k is similar to B^k.
Taking the determinant of both sides of (1) we see that $det(B) = det(A) \neq 0$ therefore if A is invertible then also B is invertible. By a simple algebraic manipulation:

$$\begin{aligned} T^{-1}AT &= B \\ AT &= TB \\ TB^{-1} &= A^{-1}T \\ B^{-1} &= T^{-1}A^{-1}T \end{aligned}$$

so A^{-1} and B^{-1} are also similar.

4 Franklin Chapter 4, Problem 4, page 83.
Suppose $\det A = 0$. Show that, for all sufficiently small $|\epsilon| > 0$, $\det(A + \epsilon I) \neq 0$.

By expansion of determinant we have

$$\det(A + \epsilon I) = \epsilon^n + c_{n-1}\epsilon^{n-1} + \ldots + c_0 \equiv P(\epsilon)$$

a polynomial in ϵ. We know that $P(0) = \det(A) = 0$. So we wish to prove that for all $\delta > 0$ we can find an $\epsilon \in [-\delta, \delta]$ $(\epsilon \neq 0)$ such that $P(\epsilon) \neq 0$. We do this by contradiction, that is, suppose that for all $\delta > 0$ we can find an $\epsilon \in [-\delta, \delta]$ $(\epsilon \neq 0)$ such that $P(\epsilon) = 0$. So then, given $\delta_1 > 0$ there exists $\epsilon_1 \in [-\delta_1, \delta_1]$, $\epsilon_1 \neq 0$ with $P(\epsilon_1) = 0$. Take $\delta_2 = \frac{\epsilon_1}{2}$, then there exists $\epsilon_2 \in [-\delta_2, \delta_2]$, $\epsilon_2 \neq 0$ with $P(\epsilon_2) = 0$ Then take $\delta_2 = \frac{\epsilon_1}{2}$, ... Repeat the process $m > n$ times so that we have found m zeros of P within the interval $[-\delta, \delta]$. However, since P is a polynomial of degree n with $m > n$ roots it must be identically zero. This is a contradiction since we know at the very least that the coefficient of ϵ^n is nonzero. Hence we must have that for all sufficiently small $|\epsilon| > 0$, $\det(A + \epsilon I) \neq 0$.

5 Let u be a vector in R^n with coordinates $(u_1, u_2, \ldots, u_n)$ and let P be the orthogonal projection onto the line $\{\lambda u, \lambda \in R\}$.

1. Show that this is a linear transformation.
2. Find the matrix of P in the canonical basis. We will also call this matrix P. (The canonical basis $e^1, \ldots, e^n$ is the basis where e^i is the vector with a one on the ith position and zeros everywhere else.)
3. What is P^2?
4. What is the trace of P? What is the rank of P?
5. Repeat 1., 3., 4., where P is now the orthogonal projection onto r linearly independent vectors. What is P in the case $r = n$?

Let K denote the space$\{\lambda u, \lambda \in R\}$. P is a tranformation that takes $x \in R^n$ to $x^K = Px = \lambda u$, such that $\lambda = \min_{\lambda \in R} \parallel \lambda u - x \parallel$ We can find a formula for Px. For a fixed $x \in R^n$, if we onsider the geometrical interpretation of a projection operator, it is easy to be convinced that $\parallel \lambda u - x \parallel$ achieves a minimum for some λ. For the same λ, $\parallel \lambda u - x \parallel^2$ is also minimized.
Then it must be:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \parallel \lambda u - x \parallel^2 &= \frac{\partial}{\partial \lambda}(\lambda u_1 - x_1)^2 + \ldots + (\lambda u_n - x_n)^2 = 0 \\ &= (\lambda u_1 - x_1)u_1 + \ldots + (\lambda u_n - x_n)u_n = 0 \\ \Rightarrow \lambda &= \frac{x_1 u_1 + \ldots + x_n u_n}{u_1 u_1 + \ldots + u_n u_n} = \frac{x \cdot u}{u \cdot u} \end{aligned}$$

1. Let $\alpha \in R$, $Px^1 = \lambda_1 u$, $Px^2 = \lambda_2 u$ $Px = \lambda u$:

$$P(\alpha x) = \frac{(\alpha x) \cdot u}{u \cdot u} u = \alpha \frac{x \cdot u}{u \cdot u} u = \alpha P x$$

and, if $x = x^1 + x^2$

$$P(x^1 + x^2) = \frac{(x^1 + x^2) \cdot u}{u \cdot u} u = \frac{x^1 \cdot u}{u \cdot u} u + \frac{x^2 \cdot u}{u \cdot u} u = Px^1 + Px^2$$

This proves linearity of the projection operator

2. Using the linearity of the operator, we have:

$$Px = P(x_1 e^1 + \ldots + x_n e^n) = x_1 Pe^1 + \ldots + x_n Pe^n$$

Note that $Pe^i = \frac{e^i \cdot u}{u \cdot u} u = \frac{u_i}{u \cdot u} u$, so the matrix P will be

$$P = \begin{pmatrix} \frac{u_1}{u \cdot u} u_1 & \frac{u_2}{u \cdot u} u_1 & \ldots & \frac{u_n}{u \cdot u} u_1 \\ \frac{u_1}{u \cdot u} u_2 & \frac{u_2}{u \cdot u} u_2 & \ldots & \frac{u_n}{u \cdot u} u_2 \\ \cdot & \cdot & \ldots & \cdot \\ \frac{u_1}{u \cdot u} u_n & \frac{u_2}{u \cdot u} u_n & \ldots & \frac{u_n}{u \cdot u} u_n \end{pmatrix}$$

3. P^2 acting on x, is the same as P acting on Px. Since Px is already a vector in the subspace K, applying P again to it does not change anything. So we must have:

$$P^2 = P$$

4. The trace is the sum of the diagonal elements. In this case it is:

$$\mathrm{tr}(A) = \frac{u_1}{u \cdot u} u_1 + \frac{u_2}{u \cdot u} u_2 + \ldots + \frac{u_n}{u \cdot u} u_n = \frac{u \cdot u}{u \cdot u} = 1$$

The rank is the dimension of the column space (range of the matrix). The column space is the vector u and thus $\mathrm{rank}(A) = 1$.

5. Now let K denote the subspace spanned by the linearly independent vectors $u^1, \ldots, u^r$. P is a tranformation that takes $x \in R^n$ to $x^K = Px = \lambda_1 u^1 + \ldots + \lambda_r u^r$ in K, such that $\parallel x^K - x \parallel$ is minimized for $\lambda_1, \ldots, \lambda_r \in R$. Define λ to be the r dimensional vector $(\lambda_1, \ldots, \lambda_r)$ and A to be the $n \times r$ matrix whose columns are the vectors $u^1, \ldots, u^r$. Then we are trying to find $\min_{\lambda \in R^r} \parallel A\lambda - x \parallel$. This is the solution of

$$A^T A\lambda = A^T x. \tag{3}$$

Equation (3) is linear and the linearity of P follows from this fact. Let $\alpha \neq 0 \in R$, $\lambda_1, \lambda_2, \lambda \in R^r$, $Px^1 = A\lambda_1$, $Px^2 = A\lambda_2$, $Px = A\lambda$:
If x, $A\lambda$ satisfy (3), then αx, $\alpha A\lambda$ also satisfy (3):

$$\begin{array}{rcl} & P(\alpha x) & \\ A^T \overbrace{\alpha A\lambda} & = & A^T(\alpha x) \\ A^T \alpha P(x) & = & A^T(\alpha x) \\ \Rightarrow P(\alpha x) & = & \alpha P(x) \end{array}$$

Let $\lambda = \lambda_1 + \lambda_2$. If x^1, $A\lambda_1$ and x^2, $A\lambda_2$ satisfy (3), then $x^1 + x^2$, $A(\lambda_1 + \lambda_2)$ also satisy (3):

$$\begin{array}{rcl} & P(x^1+x^2) & \\ A^T\overbrace{A(\lambda_1+\lambda_2)} & = & A^T(x^1+x^2) \\ A^T(Px^1+Px^2) & = & A^T(x^1+x^2) \\ \Rightarrow P(x^1+x^2) & = & Px^1+Px^2 \end{array}$$

Using the same argument as for the case when P is the projection onto a line, we can get $P^2 = P$.

The column space is of course the span of our r independent vectors and therefore $\text{rank}(P) = r$.

It is a bit more tricky to find the trace. To calculate the trace we will represent the orthogonal projection in a convenient basis and then use the fact that the trace is an invariant. Let $(u_1, u_2, \ldots, u_r, u'_1, u'_2, \ldots u'_{n-r}$ be the basis of R^n obtained by merging the u's with $n-r$ independent and orthogonal vectors. In that basis, the matrix P is the matrix with ones on the first r diagonal elements and zeroes everywhere else. The trace is then the number of ones which is r. Hence, the trace of P is equal to r since it is an invariant.

If $r = n$, then $K = R^n$. The vector Px in R^n that minimizes $\parallel Px - x \parallel$ is the vector x itself so $P = I$ for $r = n$.

6 Let A be an arbitrary matrix. Show that A is not similar to $A + I$.

Suppose A were similar to $A + I$. Then:

$$\text{tr}(A) = \text{tr}(A + I)$$

Note that for any two matrices U, V we have $\text{tr}(U + V) = \text{tr}(U) + \text{tr}(V)$. Then

$$\begin{array}{rcl} \text{tr}(A) & = & \text{tr}(A) + \text{tr}(I) \\ 0 & = & \text{tr}(I) \end{array}$$

which is a contradiction. Therefore A cannot be similar to $A + I$.

7 Consider any A and a special "plane rotation" M.

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, M = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Choose the rotation θ as to annihilate the $(3,1)$ entry of $M^{-1}AM$.

M^{-1} is the rotation by an angle $-\theta$. Then:

$$\begin{array}{rcl} M^{-1}AM & = & \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ & = & \begin{pmatrix} a\cos\theta + d\sin\theta & b\cos\theta + e\sin\theta & c\cos\theta + f\sin\theta \\ -a\sin\theta + d\cos\theta & -b\sin\theta + e\cos\theta & -c\sin\theta + f\cos\theta \\ g & h & i \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

The (3,1) entry of $M^{-1}AM$ is $g\cos\theta + h\sin\theta$. Setting it to zero, we get $\tan\theta = -\frac{g}{h}$, if $h \neq 0$ and $\theta = \frac{\pi}{2}$, if $h = 0$.