18.06 Final Solution

Hold on Tuesday, 19 May 2009 at 9am in Walker Gym.
Total: 100 points.

Problem 1:

A sequence of numbers $f_0, f_1, f_2, \ldots$ is defined by the recurrence

$$f_{k+2} = 3f_{k+1} - f_k,$$

with starting values $f_0 = 1$, $f_1 = 1$. (Thus, the first few terms in the sequence are $1, 1, 2, 5, 13, 34, 89, \ldots$.)

(a) Defining $\mathbf{u}_k = \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix}$, re-express the above recurrence as $\mathbf{u}_{k+1} = A\mathbf{u}_k$, and give the matrix A.

(b) Find the eigenvalues of A, and use these to predict what the ratio f_{k+1}/f_k of successive terms in the sequence will approach for large k.

(c) The sequence above starts with $f_0 = f_1 = 1$, and $|f_k|$ grows rapidly with k. Keep $f_0 = 1$, but give a *different* value of f_1 that will make the sequence (with the *same recurrence* $f_{k+2} = 3f_{k+1} - f_k$) approach *zero* ($f_k \to 0$) for large k.

Solution (18 points = 6+6+6)

(a) We have

$$\begin{pmatrix} f_{k+2} \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix} \quad \Rightarrow \quad A = \begin{pmatrix} 3 & -1 \\ 1 & 0 \end{pmatrix}.$$

(b) Eigenvalues of A are roots of $\det(A - \lambda I) = \lambda^2 - 3\lambda + 1 = 0$. They are $\lambda_1 = \dfrac{3+\sqrt{5}}{2}$ and $\lambda_2 = \dfrac{3-\sqrt{5}}{2}$. Note that $\lambda_1 > \lambda_2$, so the ratio f_{k+1}/f_k will approach $\lambda_1 = \dfrac{3+\sqrt{5}}{2}$ for large k.

(c) Let $\mathbf{v}_1, \mathbf{v}_2$ be the eigenvectors with eigenvalues λ_1 and λ_2, respectively. So, we can write $\mathbf{u}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ and then $\mathbf{u}_k = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2$. If we need $f_k \to 0$, we have to make $c_1 = 0$. In other words, $\mathbf{u}_0$ must be proportional to the eigenvector $\mathbf{v}_2$.

$$A - \lambda_2 I = \begin{pmatrix} \frac{3+\sqrt{5}}{2} & -1 \\ 1 & -\frac{3-\sqrt{5}}{2} \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_2 = \begin{pmatrix} \frac{3-\sqrt{5}}{2} \\ 1 \end{pmatrix}.$$

Hence, we need to take $f_1 = \dfrac{3-\sqrt{5}}{2}$ so that f_k will approach zero for large k.

Problem 2: For the matrix $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}$ with rank 2, consider the system of equations $A\mathbf{x} = \mathbf{b}$.

(i) $A\mathbf{x} = \mathbf{b}$ has a solution whenever $\mathbf{b}$ is orthogonal to some nonzero vector $\mathbf{c}$. Explicitly compute such a vector $\mathbf{c}$. Your answer can be multiplied by any overall constant, because $\mathbf{c}$ is any basis for the __________________ space of A.

(ii) Find the orthogonal projection $\mathbf{p}$ of the vector $\mathbf{b} = \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix}$ onto $C(A)$. (*Note*: The matrix $A^{\mathrm{T}}A$ is singular, so you *cannot* use your formula $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ to obtain the projection matrix P onto the column space of A. But I have repeatedly discouraged you from computing P explicitly, so you don't need to be reminded anyway, right?)

(iii) If $\mathbf{p}$ is your answer from (ii), then a solution $\mathbf{y}$ of $A\mathbf{y} = \mathbf{p}$ minimizes what? [You need not answer (ii) or compute $\mathbf{y}$ for this part.]

Solution (18 points = 7+7+4)

(i) The system of equations $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ lies in the column space of A, which is orthogonal to the left nullspace of A. We solve for a (nonzero) vector $\mathbf{c}$ in the left nullspace using Gaussian elimination, as follows.

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{c} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

The answer can by any nonzero multiple of $\mathbf{c}$, which will be a basis for the *left nullspace* of A.

(ii) Method 1: Since $\mathbf{c}$ is a basis of the orthogonal complement of the column space $C(A)$, the projection of $\mathbf{b}$ onto $C(A)$ can be computed as

$$\mathbf{p} = \mathbf{b} - \frac{\mathbf{c}^T\mathbf{b}\mathbf{c}}{\|\mathbf{c}\|^2} = \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix} - \frac{-9}{3}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 12 \end{pmatrix}.$$

Method 2: (not recommended) We know that $\mathbf{p}$ is the best linear approximation of $\mathbf{b}$. So we solve

$$A^{\mathrm{T}}A\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = A^{\mathrm{T}}\begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix},$$

$$\begin{pmatrix} 6 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 36 \\ 18 \\ 0 \end{pmatrix}$$

We can get a particular solution $\mathbf{y} = (6, 0, 0)^{\mathrm{T}}$. (There are other solutions too.) Hence,

$$\mathbf{p} = A\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 12 \end{pmatrix}.$$

(iii) Since $\mathbf{p}$ is the orthogonal projection of $\mathbf{b}$ onto $C(A)$, A solution $\mathbf{y}$ of $A\mathbf{y} = \mathbf{p}$ minimizes the distance $\|A\mathbf{y} - \mathbf{b}\|$.

Problem 3: True or false. Give a counter-example if *false*. (You need not provide a reason if true.)

(a) If Q is an orthogonal matrix, then $\det Q = 1$.

(b) If A is a Markov matrix, then $d\mathbf{u}/dt = A\mathbf{u}$ approaches some finite constant vector (a "steady state") for any initial condition $\mathbf{u}(0)$.

(c) If S and T are subspaces of $\mathbb{R}^2$, then their intersection (points in *both* S *and* T) is also a subspace.

(d) If S and T are subspaces of $\mathbb{R}^2$, then their union (points in *either* S *or* T) is also a subspace.

(e) The rank of AB is less than or equal to the ranks of A and B for any A and B.

(f) The rank of $A + B$ is less than or equal to the ranks of A and B for any A and B.

Solution (12 points = 2+2+2+2+2+2)

(a) False. For example, $Q = (-1)$ is an orthogonal matrix: $Q^{\mathrm{T}}Q = (-1)(-1) = (1)$.

REMARK: In general, for a real orthogonal matrix Q, $\det Q = \pm 1$. This is because $\det(Q^{\mathrm{T}}Q) = \det(I) = 1 \;\Rightarrow\; \det(Q)^2 = \det(Q^{\mathrm{T}})\det(Q) = 1$.

(b) False. Be careful here that we are discussing differential equations but not the powers of A. For example, $A = (1)$, the differential equation has solution $\mathbf{u} = ce^t$ for some constant c, which does not approach to any finite constant vector.

REMARK: It is true that for the Markov process $\mathbf{u}_{k+1} = A\mathbf{u}_k$, $\mathbf{u}_k$ approaches some finite constant vector (a "steady state") for any initial condition $\mathbf{u}_0$.

(c) True. Intersections of subspaces are always subspaces.

(d) False. For example, S and T are the x- and y-axes. Then $(1,1) = (1,0) + (0,1)$ is a linear combination of points in the union of S and T, but does not lie in the union itself. So the union of S and T is not a subspace.

(e) Ture. One may see this by arguing as follows. Since the column space of AB is a subspace of the column space of A, the rank of AB is less than or equal to the rank of A. Similarly, since the row space of AB is a subspace of the row space of B, the rank of AB is less than or equal to the rank of B.

(f) False. $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ both have rank 1. But $A + B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has rank 2.

REMARK: It is true that $\text{rank}(A + B) \leq \text{rank}A + \text{rank}B$.

Problem 4: Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 0 & -3 \\ 1 & 0 & -1 \end{pmatrix}$$

(a) Find an orthonormal basis for $C(A)$ using Gram-Schmidt, forming the columns of a matrix Q.

(b) Write each step of your Gram-Schmidt process from (a) as a multiplication of A on the ________ (*left* or *right*) by some invertible matrix. Explain how the product of these invertible matrices relates to the matrix R from the QR factorization $A = QR$ of A.

(c) Gram-Schmidt on another matrix B (of the same size as A) gives the *same* orthonormal basis (the same Q) as in part (a). Which of the four subspaces, if any, must be the same for the matrices AA^{T} and BB^{T}? [*You can do this part without doing (a) or (b).*]

Solution (18 points = 6+6+6)

(a) From $\mathbf{u}_1 = (1,1,1,1)^{\mathrm{T}}$, we get $\mathbf{q}_1 = \mathbf{u}_1/\|\mathbf{u}_1\| = \frac{1}{2}(1,1,1,1)^{\mathrm{T}}$.

$$\begin{aligned}
\mathbf{v}_2 &= (1,-1,0,0)^{\mathrm{T}}, \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathbf{q}_1^{\mathrm{T}}\mathbf{v}_2\mathbf{q}_1 = \mathbf{v}_2 = (1,-1,0,0)^{\mathrm{T}}, \\
\mathbf{q}_2 &= \mathbf{v}_2/\|\mathbf{v}_2\| = \frac{1}{\sqrt{2}}(1,-1,0,0)^{\mathrm{T}}; \\
\mathbf{v}_3 &= (1,-1,-3,-1)^{\mathrm{T}}, \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathbf{q}_1^{\mathrm{T}}\mathbf{v}_3\mathbf{q}_1 - \mathbf{q}_1^{\mathrm{T}}\mathbf{v}_3\mathbf{q}_1 = \mathbf{v}_3 + \mathbf{u}_1 - \mathbf{u}_2 = (1,1,-2,0)^{\mathrm{T}}, \\
\mathbf{q}_3 &= \mathbf{v}_3/\|\mathbf{v}_3\| = \frac{1}{\sqrt{6}}(1,1,-2,0)^{\mathrm{T}}.
\end{aligned}$$

Hence, we have

$$Q = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/2 & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/2 & 0 & -2/\sqrt{6} \\ 1/2 & 0 & 0 \end{pmatrix}.$$

(b) Each step of the Gram Schmidt process from (a) is a multiplication of A on the *right* as follows.

$$\begin{aligned}
A \rightsquigarrow A \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &\rightsquigarrow A \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&\rightsquigarrow A \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \\
&\rightsquigarrow A \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/\sqrt{6} \end{pmatrix} = Q.
\end{aligned}$$

The product of these invertible 3×3 matrices is exactly R^{-1}.

(c) Since the Gram-Schmidt of A and B gives the same outcome, the column space of A and B are the same. We know that A and AA^{T} have the same column space, and B and BB^{T} have the same column space. Hence AA^{T} and BB^{T} have the same column space. Moreover, since left nullspace is always orthogonal to the column space, AA^{T} and BB^{T} have the same left nullspace too. Also, notice that AA^{T} and BB^{T} are symmetric matrices, their row spaces are the same as the column spaces, and their nullspaces are the same as the left nullspaces. Therefore, all four subspaces of AA^{T} are the same as BB^{T}.

Problem 5: The complete solution to $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

for any arbitrary constants c and d.

(i) If A is an $m \times n$ matrix with rank r, give as much true information as possible about the integers m, n, and r.

(ii) Construct an explicit example of a possible matrix A and a possible right-hand side $\mathbf{b}$ with the solution $\mathbf{x}$ above. (There are many acceptable answers; you only have to provide one.)

Solution (16 points = 8+8)

(i) Since we can multiply A with $\mathbf{x}$, $n = 3$. Also, since the nullspace of A is 2-dimensional, $r = n - 2 = 1$. There is no restriction on m except that $m \geq r = 1$.

(ii) We construct a minimal one, namely, $A = (a_1\ a_2\ a_2)$ is 1×3. For this, we need $A \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 0$ and $A \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = 0$. That is

$$\begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

A special solution is $A = (1\ -1\ 2)$. In this case, $\mathbf{b} = A\mathbf{x} = (1\ -1\ 2) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (-1)$.

So, an example is

$$\begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \mathbf{x} = (-1).$$

Problem 6: Consider the matrix

$$A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

(i) A has one eigenvalue $\lambda = -1$, and the other eigenvalue is a double root of $\det(A - \lambda I)$. What is the other eigenvalue? (Very little calculation required.)

(ii) Is A defective? Why or why not?

(iii) Using the above A, suppose we want to solve the equation

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} + c\mathbf{u}$$

where c is some real number, for some initial condition $\mathbf{u}(0)$.

(a) For what values of c will the solutions $\mathbf{u}(t)$ always to go zero as $t \to \infty$?

(b) For what values of c will the solutions $\mathbf{u}(t)$ typically diverge ($\|\mathbf{u}(t)\| \to \infty$) as $t \to \infty$?

(c) For what values of c will the solutions $\mathbf{u}(t)$ approach a constant vector (possibly zero) as $t \to \infty$?

Solution (18 points = 6+6+6 (2+2+2))

(i) Let $\lambda_1 = -1$ and let $\lambda_2 = \lambda_3$ denote the double roots. Then from the trace of A, we have $\lambda_1 + 2\lambda_2 = \text{trace}(A) = 3$. Hence, $\lambda_2 = 2$.

(ii) A is not defective. There are two ways to see it. For one way, since A is symmetric, it is always non-defective; for another way, we compute $A - \lambda_2 I = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}$, which has rank 1 and hence its nullspace is 2-dimentional.

(iii) The key point here is that $A + cI$ would have eigenvalues $\lambda_1 + c$ and $\lambda_2 + c$ (with multiplicity 2). An alternative point of view is as follows. If we write the initial condition $\mathbf{u}(t) = c_1(t)\mathbf{v}_1 + c_2(t)\mathbf{v}_2 + c_3(t)\mathbf{v}_3$, then the differential equation becomes

$$\frac{dc_1(t)}{dt}\mathbf{v}_1 + \frac{dc_2(t)}{dt}\mathbf{v}_2 + \frac{dc_3(t)}{dt}\mathbf{v}_3 = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + c_3\lambda_3\mathbf{v}_3 + cc_1\mathbf{v}_1 + cc_2\mathbf{v}_2 + cc_3\mathbf{v}_3.$$

We have

$$\begin{cases} \dfrac{dc_1(t)}{dt} = c_1\lambda_1 + cc_1, & \Rightarrow \quad c_1 = e^{(\lambda_1+c)t}; \\ \dfrac{dc_2(t)}{dt} = c_2\lambda_2 + cc_2, & \Rightarrow \quad c_2 = e^{(\lambda_2+c)t}; \\ \dfrac{dc_3(t)}{dt} = c_3\lambda_3 + cc_3, & \Rightarrow \quad c_3 = e^{(\lambda_3+c)t}; \end{cases}$$

(a) If we require $\mathbf{u}(t)$ always go zero as $t \to \infty$, $\lambda_1 + c < 0$, $\lambda_2 + c = \lambda_3 + c < 0$. Hence, we require $c < -2$.

(b) If the solution $\mathbf{u}(t)$ typically diverge, we need either $\lambda_1 + c > 0$ or $\lambda_2 + c = \lambda_3 + c > 0$. Hence, we require $c > -2$.

(c) If we allow the solution to approach to some constant vector, we allow to have the extreme case of (a), that is to say $c \leq -2$.