MIT 18.06 Final Exam Solutions, Fall 2022, Johnson

Problem 1 [5+10 points]:

$Ax = b$ has solutions $x_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $x_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, and possibly other solutions, for some (real) matrix A and right-hand side b.

(a) A is an $m \times n$ matrix with rank r. Give as **much true information as possible** about m, n, r. (For example, "$m = 16, r = 0, n \leq 12$" is a possible, but incorrect, answer.)

(b) Give another solution $x_3 =$ ________ (different from x_1 and x_2) for the same equation $Ax = b$. You can do this because you know a nonzero vector ________ in the ________ space of A.

Solution:

(a) We must have $\boxed{n = 3}$ because the solutions have 3 components. Since the solutions are not unique, A **cannot** have full column rank and so $\boxed{0 \leq r \leq 2}$. We must have $\boxed{m \geq r}$ rows (which is true for any matrix).

(b) The difference $x_2 - x_1 = \boxed{\begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}}$ between two solutions (or any multiple thereof) must be a vector in the **null space** of A. So, we can find more solutions simply by adding any multiple of this to x_1 or x_2, for example $x_2 + (x_2 - x_1) = \boxed{\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}}$ is a solution, or in fact *any* vector of the form $x_1 + \frac{\alpha}{3}(x_2 - x_1) = \boxed{\begin{pmatrix} \alpha + 1 \\ \alpha + 2 \\ \alpha + 3 \end{pmatrix}}$ for any scalar α (this is the "complete" solution to $Ax = b$, though you weren't required to write this explicitly).

Problem 2 [10+5 points]:

Robert "Bobby Boy" Boyle (way back in 1662) measured a sequence of m data points $(p_1, v_1), (p_2, v_2), \ldots, (p_m, v_m)$ relating the pressure p of a gas to its volume v. Suppose that he wanted to fit his data to a model of the form

$$V(P) = \alpha + \frac{\beta}{P}$$

and solve for the unknown coefficients α and β that minimize the sum-of-squares error $\sum_k [v_k - V(p_k)]^2$ between the model and the measured data.

(a) Write down a ____×____ system of linear equations (matrix?)(unknowns?) = (right-hand side?) that Bobby could solve to find these best-fit coefficients α and β. You can leave the matrix and right-hand-side as products of terms involving other matrices and/or vectors, but **clearly describe how each term** is constructed from the data $(p_1, v_1), (p_2, v_2), \ldots, (p_m, v_m)$.

(b) Using these best-fit α and β values, the vector $\delta = \begin{pmatrix} v_1 - V(p_1) \\ v_2 - V(p_2) \\ \vdots \\ v_m - V(p_m) \end{pmatrix}$ of discrepancies between the model and the data is an orthogonal projection of the vector ________ onto the ________ space of the matrix ________ .

Solution:

(a) There are 2 unknowns, so we will have a $\boxed{2 \times 2}$ system of equations given by the **normal equations** for our least-square problem:

$$\boxed{A^T A \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\hat{x}} = A^T b}$$

where

$$\boxed{A = \underbrace{\begin{pmatrix} 1 & \frac{1}{p_1} \\ 1 & \frac{1}{p_2} \\ \vdots & \vdots \\ 1 & \frac{1}{p_m} \end{pmatrix}}_{m \times 2}, \qquad b = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}}$$

since we want to minimize $\sum_k [v_k - V(p_k)]^2 = \|b - Ax\|^2$.

(b) The vector δ is precisely the error ("residual") $\delta = b - A\hat{x}$. Recall that the least-square solution $\hat{x}$ is chosen so that $p = A\hat{x} = Pb$ is the projection of b onto $C(A)$, and $\delta = b - A\hat{x} = b - p = (I - P)b$ is the **projection of** $\boxed{b}$

onto $\boxed{N(A^T)}$**,the left nullspace of** $\boxed{A}$.

(If you've forgotten this, it's always useful to draw a sketch of least-square fitting to remind yourself that the $b - Ax$ is minimized when Ax is the orthogonal projection of b onto $C(A)$.)

Problem 3 [5+10 points]:

Consider the system of differential equations

$$\frac{dx}{dt} = \begin{pmatrix} -1 & 2 \\ & a \end{pmatrix} x$$

with initial condition $x(0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

(a) For what value(s) of a will the solution $x(t)$ approach a nonzero constant vector at large t?

(b) Using the value of a from the previous part, write down the exact solution $x(t)$ (at all times, not just for large t).

Solution:

(a) To make the ODE solution $x(t) = e^{At}$ go to a nonzero constant, we want one $e^{\lambda t}$ term to be constant (i.e. $\lambda = 0$) and the other $e^{\lambda t}$ term to be decaying (i.e. $\text{Re}(\lambda) < 0$). Since $A = \begin{pmatrix} -1 & 2 \\ & a \end{pmatrix}$ is an upper-triangular, $\det(A - \lambda I)$ is just the product of the diagonals $(-1-\lambda)(a-\lambda)$ and the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = a$. The first eigenvalue gives decaying solutions, so we need $\boxed{a = 0}$ to get a constant solution from the other term.

Technically, we also need to check that $x(0)$ has a nonzero coefficient of the $\lambda_2 = 0$ eigenvector, but we will verify this in part (b).

(b) To obtain $x(t)$, we need to (1) expand $x(0)$ in the basis of eigenvectors and (2) multiply each term by $e^{\lambda t}$. That is, we are looking for the solution:

$$x(t) = e^{At}x(0) = \underbrace{\begin{pmatrix} x_1 & x_2 \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} e^{-t} & \\ & e^{0t} \end{pmatrix}}_{e^{\Lambda t}} \underbrace{X^{-1}x(0)}_{c} = c_1 e^{-t} x_1 + c_2 x_2,$$

which corresponds to expanding a solution in the basis of the eigenvectors x_1, x_2, finding the coefficients c from $x(0)$, and multiplying each term by the corresponding $e^{\lambda t}$.

First, we need to *find* the eigenvectors, but this a straightforward exercise in computing nullspaces (which in this simple case can be done by inspection):

$$(A - \overset{-1}{\cancel{\lambda_1}} I)x_1 = \begin{pmatrix} 0 & 2 \\ & 1 \end{pmatrix} x_1 = \vec{0} \implies x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$(A - \cancelto{0}{\lambda_2} I)x_2 = \begin{pmatrix} -1 & 2 \\ & 0 \end{pmatrix} x_2 = \vec{0} \implies x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Now, to expand $x(0)$ this basis, we write

$$x(0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = c_1 \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{x_1} + c_2 \underbrace{\begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{x_2} = \underbrace{\begin{pmatrix} 1 & 2 \\ & 1 \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{c} \implies c = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which is solvable by inspection, or by using the fact that X is upper-triangular so we can do backsubstitution (with no elimination steps). (If we wrote the eigenvalues in the opposite order we would have gotten a lower-triangular X, from which we could do forward-substitution.) Hence

$$x(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 = e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{0t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \boxed{\begin{pmatrix} 2 + e^{-t} \\ 1 \end{pmatrix}},$$

which clearly approaches the nonzero constant vector x_2 as desired in part (a), since $c_2 \neq 0$.

Problem 4 [4+4+4+4+4 points]:

The following short-answer questions are answered independently (and refer to unrelated matrices A for each part), requiring little or no computation:

(a) Any solution x of $Ax = b$ is a sum of a vector in the ________ space of A and a vector in the in the ________ space of A.

(b) If $Ax = b$ is solvable for *any* b, then it might be a (**circle one**) 10×3 or 3×10 matrix with rank $r =$ ________. If $Ax = b$ has a *unique* solution x for some b then it might be a (**circle one**) 10×3 or 3×10 matrix with rank $r =$ ________.

(c) Relate the four fundamental subspaces of $A^T A$ to the four fundamental subspaces of a real matrix A: nullspace of $A^T A =$ ________ space of A, left nullspace of $A^T A =$ ________ space of A, column space of $A^T A =$ ________ space of A, row space of $A^T A =$ ________ space of A.

(d) Suppose we solve $A^T A\hat{x} = A^T b$ for $\hat{x}$ given some real A. Then, the orthogonal projection of b into $C(A)$ is the vector ________ and the projection of b onto $N(A^T)$ is the vector ________. (Give formulas in terms of $A, b, \hat{x}$ involving *no matrix inverses.*)

(e) Which of the following matrices **cannot** be singular for **any** real square matrix A (circle **all** answers): $A^T A$, A^2+I, $(A+A^T)^2+I$, e^{-A}, $A+10^{100}I$, $3A^T A + 4I$.

Solution:

(a) The **row space** $\boxed{C(A^H)}$ and the **null space** $\boxed{N(A)}$, since together these give the whole space $\mathbb{R}^n$ of possible inputs of any $m \times n$ matrix A.

(b) If it's solvable for *any* b, then A must be a "wide" matrix with full row rank, for example a $\boxed{3 \times 10}$ matrix with rank $\boxed{r = 3}$. If the solutions are unique, then A must be a "tall" matrix with full column rank, for example a $\boxed{10 \times 3}$ matrix with rank $\boxed{r = 3}$.

(c) We showed in class that the nullspace of $A^T A$ matches that of A and the column space matches that of A^T. Furthermore, since $A^T A$ is real-symmetric, i.e. $(A^T A)^T = A^T A$, the same things hold true of the left nullspace and the row space. So the nullspace is $\boxed{N(A^T A) = N(A)}$, the left nullspace is $\boxed{N((A^T A)^T) = N(A)}$, the column space is $\boxed{C(A^T A) = C(A^T)}$, and the row space is $\boxed{C((A^T A)^T) = C(A^T)}$.

(d) The orthogonal projection of b onto $C(A)$ is $\boxed{A\hat{x}}$ and the projection of b onto $N(A^T)$ is $\boxed{b - A\hat{x}}$. This is how we derived the normal equations $A^T A\hat{x} = A^T b$ in the first place!

(e) $A^T A$ can be singular since it is only semidefinite (e.g. suppose $A = 0$). $A^2 + I$ can be singular if A has an eigenvalue of $\pm i$ (possible for real A!). $\boxed{(A + A^T)^2 + I}$ **cannot** be singular since $A + A^T$ is real-symmetric with real eigenvalues, so the eigenvalues of $(A + A^T)^2 + I$ are $(\text{real})^2 + 1 > 0$. $\boxed{e^{-A}}$ **cannot** be singular since $e^{-\lambda} \neq 0$ for any eigenvalue λ of A. $A + 10^{100} I$ can be singular if A has an eigenvalue $\lambda = -(10^{100})$. $\boxed{3A^T A + 4I}$ **cannot** be singular since $A^T A$ is semidefinite with eigenvalues ≥ 0, so $3A^T A + 4I$ has eigenvalues of the form $3(\text{something} \geq 0) + 4 > 0$.

Problem 5 [10+5+5 points]:

Suppose you have a matrix $A = C^{-1}B$ where

$$B = \begin{pmatrix} 1 & & \\ -1 & 2 & \\ 2 & 1 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & & 4 \\ & 2 & 2 \\ 4 & 2 & 2 \end{pmatrix}.$$

The following parts can be **answered independently**.

(a) Compute the **first column** of A^{-1}.

(b) Compute the **trace** of the matrix $A^{-1}B$. (Little calculation is required because $A^{-1}B$ has the same trace, and the same eigenvalues, as ________, since the two matrices are ________!)

(c) One of the eigenvalues of C is $\lambda_1 = 2$. A corresponding eigenvector is $x_1 =$ ________.

Solution:

(a) We can do this *without* computing A^{-1} explicitly (which is almost always a mistake). We just need to compute

$$x = A^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \left(C^{-1}B\right)^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \underbrace{B^{-1}\underbrace{C\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{b}}_{\text{triangular solve } Bx=b}.$$

The first step is $b = C\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}$, just the first column of c. The second step is to compute $x = B^{-1}b$ by solving $Bx = b$ for x, but since B is lower-triangular we can do this easily by forward-substitution:

$$\underbrace{\begin{pmatrix} 1 & & \\ -1 & 2 & \\ 2 & 1 & 1 \end{pmatrix}}_{B}\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}}_{c} \Longrightarrow \begin{matrix} x_1 = 2 \\ -x_1 + 2x_2 = 0 \Longrightarrow x_2 = 1 \\ 2x_1 + x_2 + x_3 = 4 \Longrightarrow x_3 = -1 \end{matrix} \Longrightarrow x = \boxed{\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}}.$$

Of course, there are much more laborious ways to solve this problem by explicitly inverting and multiplying a bunch of matrices.

(b) The key thing to realize is that the matrix $A^{-1}B = B^{-1}CB$ is **similar** to the matrix $\boxed{C}$, so its trace (and determinant, and eigenvalues) match those of C. By inspection, then, $\text{trace}(A^{-1}B) = \text{trace}(C) = 2+2+2 = \boxed{6}$.

Another way of seeing this is to use the "cyclic property" of the trace: $\text{trace}(A^{-1}B) = \text{trace}(\underbrace{BA^{-1}}_{=BB^{-1}C=C})$. It's not really correct terminology to say that $A^{-1}B$ and BA^{-1} are "cyclic", however—a "cyclic matrix" refers to something else entirely. But it *is* true that given a product of matrices, you can take a cyclic permutation of the product, and get the same eigenvalues as well as the same trace: (More precisely: XY and YX have identical eigenvalues for any square X and Y, and the nonzero eigenvalues are the same even for non-square X and Y! But we often don't cover this fact in 18.06.)

Actually calculating $A^{-1}B$ is a *lot* more work (even for a computer, though for matrices this tiny it hardly matters) and very error-prone (by hand), but if you managed to do it all correctly you would get $A^{-1} = \begin{pmatrix} 2 & 0 & 4 \\ 1 & 1 & 3 \\ -1 & 1 & -9 \end{pmatrix}$ and $A^{-1}B = \begin{pmatrix} 10 & 4 & 4 \\ 6 & 5 & 3 \\ -20 & -7 & -9 \end{pmatrix}$, which of course has the same trace $10 + 5 - 9 = 6$.

(c) We just need a basis for $N(C - 2I)$:

$$(C - 2I)x = \vec{0} = \begin{pmatrix} 0 & & 4 \\ & 0 & 2 \\ 4 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

but this can be done either by inspection or simply working top-to-bottom. The first two rows immediately give $x_3 = 0$ and the last row gives $4x_1 + 2x_2 = 0 \implies x_2 = -2x_1$. So, for example, we could pick $x_1 = 1$ and get an eigenvector

$$x = \boxed{\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}}$$

or any nonzero scalar multiple thereof.

Problem 6 [4+4+4+4+4 points]:

The matrix A has eigenvalues $\lambda_1 = 1$, $\lambda_2 = -2$, and $\lambda_3 = 0$, with corresponding eigenvectors $x_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $x_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$. Consider the recurrence

$$Ay_{n+1} = y_n - 3y_{n+1},$$

starting with some initial vector y_0.

(a) Give an exact formula for $y_n =$ ________ in terms of A, I, y_0, n. (For example, $y_n = (e^{nA} + 7I)y_0$ is a possible but incorrect answer.)

(b) For a typical initial vector y_0 (e.g. one chosen at random with `randn(3)` in Julia), you should expect y_n for large n to be approximately parallel to the vector ________ and **growing/decaying/oscillating/nearly constant** with n (circle **one**).

(c) Give an example of an initial vector $y_0 =$ ________ for which y_n is **decaying** towards zero with n, and for this y_0 give an *exact* numeric formula (in terms of n) for $y_n =$ ________. (There are many possible answers, but not much calculation should be needed.) Your answer should have no matrices or unknowns, only vectors of numbers or simple arithmetic expressions like 2^n or e^n or $\frac{1}{n^2}$.

(d) The matrix A **can/must/cannot be** Hermitian (circle **one**). Briefly justify your answer.

(e) For $y_0 = \begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix}$, give a good approximate formula for $y_{100} =$ ________ (numeric vector, no unknowns or matrices).

Solution:

(a) $Ay_{n+1} = y_n - 3y_{n+1} \implies Ay_{n+1} + 3y_{n+1} = (A + 3I)y_{n+1} = y_n \implies y_{n+1} = (A+3I)^{-1}y_n$. Note that $A+3I$ must be invertible because A has no eigenvalues of -3. Starting with y_0, we then get $y_1 = (A + 3I)^{-1}y_0$, followed by $y_2 = (A + 3I)^{-1}y_1 = (A + 3I)^{-2}y_0$, and so on, so

$$\boxed{y_n = (A + 3I)^{-n}y_0}$$

for any n.

(b) Since A has eigenvalues $1, -2, 0$, it follows that $(A+3I)^{-n}$ has eigenvalues $(1 + 3)^{-n}, (-2 + 3)^{-n}, (0 + 3)^{-n} = \frac{1}{4^n}, 1^n, \frac{1}{3^n}$. Two of these are decaying exponentially with n, so for large n we should expect y_n to be dominated by the 1^n term, which is parallel to $\boxed{x_2}$ and is nearly **constant** with n. (The only exception would be if the x_2 coefficient is exactly zero, which is very unlikely for a random initial vector.)

(c) To get a decaying solution, we just need y_0 to be a nonzero vector in the span of x_1 and x_3, so that the x_2 coefficient is zero. For example, we could simply pick $\boxed{y_0 = x_1}$ and get $\boxed{y_n = \frac{1}{4^n} x_1}$. More generally, we could pick $y_0 = c_1 x_1 + c_3 x_3$ for any coefficients c_1, c_3 in which case we will get $y_n = \frac{c_1}{4^n} x_1 + \frac{c_3}{3^n} x_3$.

(d) A **cannot** be Hermitian because the given eigenvectors are **not orthogonal** for distinct eigenvalues.

(e) We just need to write this initial vector in the basis of eigenvectors $y_0 = c_1 x_1 + c_2 x_2 + c_3 x_3$ and then multiply the terms by $\frac{1}{4^n}, 1^n, \frac{1}{3^n}$ respectively to get y_n. Unfortunately, since the eigenvectors are not orthogonal, we cannot simply find the coefficients by taking dot products (which would be nice because we only need the x_2 coefficient at the end), but have to solve a linear system for the coefficients:

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}}_{X = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}} \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}}_{c} = \underbrace{\begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix}}_{y_0}.$$

Proceeding by Gaussian elimination, we only need to do a single elimination step (subtract the first row of X from the third row) to get it in upper-triangular form, and the same thing to the right-hand side, yielding:

$$\underbrace{\begin{pmatrix} \boxed{1} & 1 & 1 \\ & \boxed{1} & 0 \\ & & \boxed{1} \end{pmatrix}}_{X \rightsquigarrow U} \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}}_{c} = \underbrace{\begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix}}_{y_0 \rightsquigarrow b} \implies \begin{matrix} \cdots \\ c_2 = -4 \\ \cdots \end{matrix}.$$

We don't even need to solve for c_3 and c_1 in this particular case, because U is so nice, but if we did we would easily find $c_3 = 1$ and $c_1 = 3$. So,

$$y_{100} = \frac{c_1}{4^{100}} x_1 + c_2 x_2 + \frac{c_3}{3^{100}} x_3 \approx \boxed{c_2 x_2 = -4 x_2 = \begin{pmatrix} -4 \\ -4 \\ -4 \end{pmatrix}},$$

since the x_1 and x_2 terms are negligible.

Problem 7 [5+8+5 points]:

The real Hermitian (real-symmetric) matrix A has an eigenvalue $\lambda_1 = -\frac{1}{2}$ *(clarification: with multiplicity 1, not a repeated root)* and a corresponding eigenvector $x_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$, and its other eigenvalues are all equal to 1.

(a) Give *one* example of an eigenvector of A for $\lambda_2 = 1$.

(b) The orthogonal projection of $b = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$ onto the span S of x_1 is ________ and the projection of b onto the orthogonal complement $S^\perp$ is ________.

(c) With the help of the previous part, an *exact* formula for $A^n \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \\ 2 \end{pmatrix} =$ ________ (in terms of n and explicit numerical vectors, no matrices or unknowns).

Solution:

(a) The key thing is to realize that we just need *any* nonzero vector $\perp x_1$, for example $\boxed{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}$ works.

Since A is Hermitian, any eigenvector for eigenvalues $\neq \lambda_1$ must be $\perp x_1$, i.e. in the orthogonal complement of the span of x_1, which is 4-dimensional. Since *all* of the other eigenvalues are 1, the eigenvalue of 1 must have multiplicity 4 (there are 5 eigenvalues in total, counting repeated roots) and there must be 4 eigenvectors for that eigenvalue — together with x_1, they must form a *basis* for $\mathbb{R}^5$ (since A is Hermitian therefore diagonalizable). So, the eigenvectors for $\lambda = 1$ must be the *whole* 4-dimensional subspace $\perp x_1$.

(b) The projection onto S is

$$p = \frac{x_1 x_1^T}{x_1^T x_1} b = x_1 \frac{\cancelto{7}{x_1^T b}}{\cancelto{7}{x_1^T x_1}} = \boxed{x_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}}.$$

Note that, as usual it would be equivalent but *much* more work to first compute the 5×5 rank-1 projection matrix $P = \frac{x_1 x_1^T}{x_1^T x_1} = \frac{1}{7} \begin{pmatrix} 1 & 2 & -1 & 0 & 1 \\ 2 & 4 & -2 & 0 & 2 \\ -1 & -2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 \end{pmatrix}$ and then multiply it by b. Parentheses make a big practical difference in linear algebra! The projection onto $S^\perp$ is simply

$$e = (I - P)b = b - p = b - x_1 = \boxed{\begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}.$$

Again, it would be a lot more work to first compute $I - P = \frac{1}{7} \begin{pmatrix} 6 & -2 & 1 & 0 & -1 \\ -2 & 3 & 2 & 0 & -2 \\ 1 & 2 & 6 & 0 & 1 \\ 0 & 0 & 0 & 7 & 0 \\ -1 & -2 & 1 & 0 & 6 \end{pmatrix}$ and *then* multiply it by b.

(c) We can write b as the sum of the two projections, $b = p + e$, and notice that the first term p is an eigenvector of $\lambda_1 = -\frac{1}{2}$ and the second term is in the orthogonal complement and hence (from part a) an eigenvector of $\lambda = 1$. So,

$$A^n b = \lambda_1^n p + 1^n e = \boxed{\frac{1}{(-2)^n} \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}.$$

Problem 8 [5+8+5 points]:

Suppose that Q is a 4×3 real matrix with orthonormal columns q_1, q_2, q_3.

(a) Starting from a real vector v (not in the column space of Q), give a formula for the fourth orthonormal vector q_4 that would be produced by Gram–Schmidt on q_1, q_2, q_3, v.

(b) Describe $N(Q)$, $N(Q^T)$, $N(Q^TQ)$, and $N(QQ^T)$: give the dimension and a basis for each (in terms of q_1, q_2, q_3, q_4 as needed).

(c) Suppose $b = q_1 + 2q_2 + 3q_3 + 4q_4$. Give the least-squares solution $\hat{x} =$ ________ minimizing $\|b - Qx\|$.

Solution:

(a) The Gram–Schmidt formula is

$$\boxed{q_4 = \frac{v - QQ^Tv}{\|v - QQ^Tv\|} = \frac{v - q_1q_1^Tv - q_2q_2^Tv - q_3q_3^Tv}{\left\|v - q_1q_1^Tv - q_2q_2^Tv - q_3q_3^Tv\right\|}},$$

i.e. we subtract the projection onto $C(Q)$ and then normalize. It is important that $v \notin C(Q)$ since otherwise subtracting the projection would give zero (and we would then divide by zero).

(b) Q has orthonormal columns and thus is full column rank. Hence $\boxed{N(Q) = \{\vec{0}\}} \subset \mathbb{R}^3$, i.e. it is **zero-dimensional** and the basis is the **empty set** $\{\}$ (zero basis vectors for zero dimensions). $N(Q^T) = C(Q)^\perp$ is everything perpendicular to q_1, q_2, q_3, but this is simply the **1-dimensional** space $\boxed{N(Q^T) = \text{span } q_4}$. We also know $\boxed{N(Q^TQ) = N(Q) = \{\vec{0}\}}$ and $\boxed{N(QQ^T) = N(Q^T) = \text{span } q_4}$ from the general identity $N(A^TA) = N(A)$.

(c) We want $Q\hat{x}$ to be the projection of b onto $C(Q)$, i.e. $Q\hat{x} = q_1 + 2q_2 + 3q_3$, which implies $\boxed{\hat{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}}$.