14.102 Problem Set 2 Solutions

1. Lecture Notes Exercise 105: Given an $m \times n$ matrix A, show that $S(B) \subseteq S(A)$ and $N(A') \subseteq N(B')$ whenever $B = AX$ for some matrix X. What is the geometric interpretation? *(Note: this is a repeat from last year's problem set; as such, the solution is right on the website. It is certainly worth doing, but the main reason I included it was to draw your attention to the result, which can be used to make part (e) of the next problem much less tedious.)*

   Solution: Suppose X is $n \times l$. Then B is $m \times l$. We have $S(A) = \{y \in \mathbb{R}^m | y = Ax$ for some $x \in \mathbb{R}^n\}$, and $S(B) = \{y \in \mathbb{R}^m | y = Bx$ for some $x \in \mathbb{R}^l\}$. We want to show that any $y \in S(B)$ belongs to $S(A)$ as well. We have $y = Bx = AXx = Az$, where $z = Xx, z \in \mathbb{R}^n$, implying that $y \in S(B) \Longrightarrow y \in S(A)$.

   For the second part, recall that $N(A') = \{x \in R^m | A'x = 0\}$, and $N(B') = \{x \in R^m | B'x = 0\}$. We want to show that $x \in N(A') \Rightarrow x \in N(B')$, and the proof is similar to the previous part: if $A'x = 0$, then we have $B'x = X'A'x = 0$.

   The main 'geometric' intuition to take from this is that if we construct the matrix B as AX, then each of its columns is a linear combination of the columns in A, weighted by the elements of a column of X (make sure you understand why this is true). This being the case, its columns cannot span any subspace of $\mathbb{R}^m$ which is not spanned by A. The intuition for the second part is the reverse, using the same reasoning: any vector not spanned by A cannot be spanned by B.

2. Let $A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 1 \end{pmatrix}$

   (a) Find $C = AB$

   Solution: $C = \begin{pmatrix} 1 & 6 \\ 1 & -1 \end{pmatrix}$

   (b) Find *rank* C

   Solution: 2 (its columns are linearly independent)

   (c) Find det C

   Solution: -7

   (d) Find $D = BA$

   Solution: $D = \begin{pmatrix} 1 & 3 & 0 \\ 4 & -2 & 2 \\ 1 & -4 & 1 \end{pmatrix}$

   (e) Find *rank* D (reminder: try to answer this using the result of problem 1 - without calculations)

Solution: Any two columns of D are linearly independent, so $rank D$ is at least two. But on the other hand, the result from problem 1 tells us that D is a subspace of the span of A and of B. Another way of saying this is that its rank can be no greater than either of the matrices we multiplied together to obtain it. Thus, $rank D = 2$.

(f) Find $\det D$

Solution: $\det D = 0$, since D is not full rank.

(g) Is C invertible? If so, find C^{-1}

Solution: $C^{-1} = \begin{pmatrix} \frac{1}{7} & \frac{6}{7} \\ \frac{1}{7} & -\frac{1}{7} \end{pmatrix}$

(h) Is D invertible? If so, find D^{-1}

Solution: D is not invertible, since it is not full rank.

(i) Find eigenvalues of C

Solution: We have $\begin{vmatrix} 1-\lambda & 6 \\ 1 & -1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) - 6 = \lambda^2 - 7 = 0 \Rightarrow \lambda = \pm\sqrt{7}$.

(j) Solve the following two linear systems (Hint: you will need no extra calculations!):

i. $\begin{cases} \frac{1}{7}x + \frac{6}{7}y & = & 1 \\ \frac{1}{7}x - \frac{1}{7}y & = & 0 \end{cases}$

ii. $\begin{cases} \frac{1}{7}u + \frac{6}{7}v & = & 0 \\ \frac{1}{7}u - \frac{1}{7}v & = & 1 \end{cases}$

Solution: We can collect these equations into one system as

$$\begin{aligned} \begin{pmatrix} \frac{1}{7} & \frac{6}{7} \\ \frac{1}{7} & -\frac{1}{7} \end{pmatrix} \begin{pmatrix} x & u \\ y & v \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ C^{-1} \begin{pmatrix} x & u \\ y & v \end{pmatrix} &= I \\ \begin{pmatrix} x & u \\ y & v \end{pmatrix} &= C \end{aligned}$$

This means that $x = 1, y = 1, u = 6, v = -1$.

3. Look at last year's problem set 1, #3, and its solution. It is good to understand the notion of changing bases, and of the coordinates of a vector with respect to a basis (we will use it again in discussing diagonalization). In particular, do Lecture Notes Exercise 114: what are the coordinates of $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ with respect to the following bases?

Solution. Notice that we can collect these two vectors and appeal to Lemma 112, recalling that if E is a basis for a vector space, then so is $F = EP$ for any nonsingular P, and that P then gives the coordinates of F with respect to E. In this case, we have $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = EP$, where we are given the basis E and asked to find P - but of course, P is simply the inverse of E, and its first column gives the coordinates of $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ with respect to E and its second gives the coordinates of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ with respect to E.

(a) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Solution: $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

(b) $\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$

Solution: $P = \begin{bmatrix} \frac{5}{17} & \frac{1}{17} \\ -\frac{2}{17} & \frac{3}{17} \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$

Solution: $P = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, assuming $\alpha\delta - \beta\gamma \neq 0$.

Solution: $P = \frac{1}{\alpha\delta-\beta\gamma} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}$

4. Prove Lemma 116 (note the hint in the lecture notes): Let $\{e_j\} = \{e_1, ..., e_n\}$ be a basis for $\mathbb{X}$, and let $\{b_j\} = \{b_1, ..., b_m\}$ be any set of vectors belonging to $\mathbb{X}$ with $m > n$. Then $\{b_j\}$ can not be linearly independent.

Solution: Write $E = [e_j]$ and $B = [b_j]$. Since E is a basis for $\mathbb{X}$, any b_j can be written as a linear combination of the e_j's: $b_j = \sum_{i=1}^{n} c_{ij} e_i$. Note the meaning of the subscripts here. There are n elements of $[e_j]$, and they are being summed up, weighted by n scalars $c_{1j}, ..., c_{nj}$; i indexes each of these products in the sum. j simply identifes the j^{th} element of b, and refers in c_{ij} to the j^{th} set of scalars $c_{1j}, ..., c_{nj}$; in general, there will be a different set of scalars for each b_j. Now, we can rewrite what we have so far as $B = EC$, where C is the $n \times m$ matrix collecting the scalar weights.

Now suppose the columns of B are linearly independent. Then for any $x \in \mathbb{R}^m, x \neq 0, Bx \neq 0$. Then we have

$$Bx = ECx = E\lambda \neq 0$$

where $\lambda = Cx, \lambda \in \mathbb{R}^n$. The fact that $E\lambda \neq 0$ implies that $\lambda = Cx \neq 0$. Since this must be true for all $x \neq 0$, this implies that the columns of C are linearly independent. But we know that any set of m vectors from

$\mathbb{R}^n$ must admit at least one linear independence if $m > n$, and so this is a contradiction; thus, the columns of B cannot be linearly independent.

Another way to show this would be to note that the columns of B are linearly independent only if the columns of C are linearly independent (for otherwise there would be a linear dependence among the sets of weights used to construct the columns of B from the basis E). But again, the m columns of C, which are elements of $\mathbb{R}^n$, cannot be linearly independent, and so neither can the columns of B.

5. Lecture Notes Exercise 124: Using the 'fundamental theorem of algebra' and the fact that $rank(A) = rank(A')$, show that

$$\begin{aligned} rank(A) + null(A') &= m \\ null(A) - null(A') &= n - m \end{aligned}$$

Solution: The fundamental theorem of algebra says that for any $m \times n$ matrix A,

$$rank(A) + null(A) = n$$

Similarly,

$$rank(A') + null(A') = m$$

Then the conclusions are immediate - simply replace $rank(A')$ with $rank(A)$ in the second expression to get

$$rank(A) + null(A') = m$$

and subtract the second expression from the first to get

$$rank(A) + null(A) - rank(A') - null(A') = null(A) - null(A') = n - m$$

6. Lecture Notes Exercise 129: Using the properties of transpose and inverse:

   (a) Prove that $A^{-k} = (A^k)^{-1}$
   Solution: $A^{-k} = (A^{-1})^k = A^{-1}A^{-1}\cdots A^{-1}$ (k times) $= (A^k)^{-1}$ by the property that $(AB)^{-1} = B^{-1}A^{-1}$.

   (b) Consider the matrix $Z = X(X'X)^{-1}X'$ where X is an arbitrary $m \times n$ matrix. Under what conditions on X is Z well-defined? Show that Z is symmetric. Also show that $ZZ = Z$ (i.e., that Z is **idempotent**).
   Solution: $(X'X)$ must be invertible. This can only be the case if $n \leq m$ (why?) and if $rankX = n$.
   To show that Z is symmetric, note that $(X'X)$ is symmetric, and hence so is $(X'X)^{-1}$ (why?). Now $Z' = (X(X'X)^{-1}X')' = X''(X'X)^{-1\prime}X' = X(X'X)^{-1}X' = Z$.
   To show that Z is idempotent, we check that $ZZ = (X(X'X)^{-1}X')(X(X'X)^{-1}X') = X(X'X)^{-1}(X'X)(X'X)^{-1}X' = X(X'X)^{-1}IX' = X(X'X)^{-1}X = Z$.

(Note: this is another repeat, but this one I included simply because it really is worth doing - you will use these facts, and the techniques needed to prove them, a LOT in statistics and econometrics, so it would be helpful to get them down now.)

7. Lecture Notes Exercise 150: Show that, if $[A, b]$ is singular, then and only then $X^* \neq \emptyset$, and further $\dim(X^*) = null[A, b] - 1$.

   Solution: Unfortunately, there was a typo here that I only noticed at the last second. I needed to include (for the first part) the condition that $rank([A, b]) = rank(A)$ (or I could have simply replaced $[A, b]$ singular with that).

   This is one case where we can construct a chain of equivalences to complete the proof, rather than showing one direction at a time. $X^* \neq \emptyset \Leftrightarrow b \in S(A) \Leftrightarrow rank([A, b]) = rank(A) \Leftrightarrow [A, b]$ is singular.

   Note that we can move from left to right without the added condition, but not from right to left: it is possible for $[A, b]$ to be singular but $rank([A, b]) \neq rank(A)$; more to the point, it is possible for $[A, b]$ to be singular but for X^* to nevertheless be empty. This is the case if A is itself singular and b does not lie in the span of A, so that $rank([A, b]) > rank(A)$, but b cannot be written as a linear combination of the columns of A. For example, suppose A is the identity matrix, except that the last column has all zeroes instead of a one at the bottom, and b is a column of zeroes except for its last element, which is one. Then $[A, b]$ is singular, but $b \notin S(A)$.

   For the second claim, note that the dimension of the space spanned by the solutions to the homogenous system $[A, b]y = 0$ is simply $null[A, b]$. Fixing the last coordinate of each of these solutions to be -1 describes a hyperplane of dimension one less than this space, so $\dim(X^*) = null[A, b] - 1$. We could also use the fundamental theorem of linear algebra:

$$\begin{aligned} rank([A, b]) + null([A, b]) &= n + 1 \\ null([A, b]) - 1 &= n - rank(A) = \dim(X^*) \end{aligned}$$

8. Calculate e^A for A equal to

   (a) $\begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$ (hint: diagonalize!)

   Solution: a typo here means that in fact, A is not diagonalizable. However, the hint given in the next part works: since $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, we have that $e^A = I + A = \begin{pmatrix} 3 & 1 \\ -4 & -1 \end{pmatrix}$

   Just to give an example of how to do this with a diagonalizable matrix, suppose A were $\begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}$. This yields two eigenvalues, 1 and -1, with the associated eigenvectors $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$. So we have $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $V = \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix}$, which in turn gives $V^{-1} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$. Then $e^A = \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & e^{-1} \end{bmatrix} \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

(b) $\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix}$ (hint: start with A^2, and recall that $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$)

Solution: $A^2 = \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, so $e^A = I + A + \frac{A^2}{2} = \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{pmatrix}$

9. Lecture Notes Exercise 182: Let X be an $m \times n$ matrix with $m \geq n$ and $rk(X) = n$. Show that $X'X$ is positive definite.

   Solution: $X'X$ is a symmetric $n \times n$ matrix, with rank n (this latter part follows from the result of problem 1, and implies that $X'X$ is full rank). Choose any nonzero $y \in \mathbb{R}^n$; then $y'X'Xy = (Xy)'Xy = z'z$, where $z = Xy$ is an $m \times 1$ vector. Since $z'z = \sum_{i=1}^{m} z_i^2$, we are done so long as we know that we don't have $z_i = 0$ for all i. But this cannot be the case; if $z = Xy = 0$, then we have $X'Xy = 0$, but because $X'X$ is full rank, for nonzero y, $X'Xy \neq 0$. Thus, $y'X'Xy > 0$ for all nonzero y, which is the definition of positive definiteness.

10. Lecture Notes Exercise 183: Show that a positive definite matrix is nonsingular.

    Solution: We know that such a matrix has all eigenvalues strictly positive; moreover, we know that its determinant is the product of these eigenvalues, which can thus not be zero, and so the matrix is nonsingular.

    (Conclude from the past two exercises that so long as $m \geq n$ and $rk(X) = n$ - as you will generally assume when you estimate systems of equations - that you don't need to wonder whether the term $(X'X)^{-1}$ is defined.)