18.06 Problem Set 10 Solutions

1. Do problem 5 from section 6.5.

   Solution $f = (x+3y)(x+y) = (x+2y+y)(x+2y-y) = (x+2y)^2 - y^2$. There are many points where this is negative, say $(-1, 2)$, where the above is $0^2 - 2^2 = -4$.

   This goes to show that not everything is positive-definite, even if all the entries are positive.

2. Do problem 26 from section 6.5.

   Solution Let's do an LU of $A = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{pmatrix}$. We should immediately get $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix}$.

   By symmetry, we don't have to do much more. We know $D = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ from the pivots, so to get C^T we should multiply L by $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ to get $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}$, and $C = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$.

   The second matrix is similar. LU gives $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{pmatrix}$. We know the square roots of D are $1, 1, \sqrt{5}$, so $C^T = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & \sqrt{5} \end{pmatrix}$ and $C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{5} \end{pmatrix}$.

   The point of this is just that the cholesky decomposition really is just LU for a symmetric matrix - don't need to think of them as separate things.

3. Do problem 6 from section 6.7.

   Solution $A^TA = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ and $AA^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Because we know that they "basically" have the same eigenvalues I'm going to save work by using the eigenvalues of AA^T, which are 3 and 1 (so we know A^TA has eigenvalues $3, 1, 0$).

   In the order of $3, 1, 0$, the normalized eigenvectors of A^TA are $(1/\sqrt{6}, 2\sqrt{6}, 1\sqrt{6}), (1/\sqrt{2}, 0, -1/\sqrt{2}), (1/\sqrt{6}, -1\sqrt{6}, 1\sqrt{6})$ and the normalized eigenvectors of AA^T are $(1/\sqrt{2}, 1/\sqrt{2}), (1/\sqrt{2}, -1/\sqrt{2})$. Multiplying, we get

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix},$$

   which we can check to be correct.

   Okay, I just realized you guys already did this in the last pset. Ugh.

4. Do problem 11 from section 6.7.

   Solution

Orthogonality tells us that $A^T A$ is going to be diagonal with entries $\rho_1^2, \ldots, \rho_n^2$. Thus, the columns of V (or the rows of V^T) are just going to be the eigenvectors $e_i = (0, \ldots, 1, \ldots, 0)$, so $V = V^T$ is going to be the n by n identity matrix.

Σ is going to have the lengths on the diagonal, because those are exactly the positive square roots of the eigenvalues.

Finally, $u_j = Av_j/\rho_j$, which in our case is exactly the normalized column $w_j/|w_j|$. So U is just going to have columns of A, but normalized.

5. Do problem 13 from section 6.7.

   Solution Let's go through the process for the SVD of R. note that $R^T R = R^T Q^T Q R = A^T A$, so the eigenvalues (and thus Σ) and eigenvectors (and thus V) remain the same in the two calculations - the only thing that changes is U.

   An alternate way to see this is to note that if we multiply U on the left by a n orthonormal Q, the result QU is still orthonormal because $(QU)^T QU = U^T Q^T QU = I$. Thus, since $R = USV^T$, $A = QR = (QU)SV^T$ is a valid SVD for A.

6. Do problem 2 from section 8.1.

   Solution Here we have $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$. $A_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, so the inverse is just $A_1^{-1} C_1^{-1} (A_1^T)^{-1} = \begin{pmatrix} 1/c_1 & 1/c_1 & 1/c_1 \\ 1/c_1 & 1/c_1 + 1/c_2 & 1/c_1 + 1/c_2 \\ 1/c_1 & 1/c_1 + 1/c_2 & 1/c_1 + 1/c_2 + 1/c_3 \end{pmatrix}$.

7. Do problem 5 from section 8.1.

   Solution The solution of this is $y = -\int f(x) + C$, with C determined by $y(1) = 0$. For $f(x) = 1$ we get $y = -x + 1$.

8. Do problem 6 from section 7.1.

   Solution Let's call the conditions "additivity" and "scaling" respectively.

   [a]: This is scaling the vector into a normal vector. Thus it is impossible that we get additivity, because the sums of normal vectors don't have to be normal. Take $T(0, 1)$ and $T(1, 0)$ for instance. Howver, true to its name this does have the scaling property, as whatever c we introduce will be canceled from v and $||v||$.

   [b]: This satisfies both. One immediate way to see this is to see that this is exactly matrix multiplication by $[1, 1, 1]$, which is a linear operation and thus satisfies both properties.

   [c]: This also satisfies both. Again, this is jsut because this is matrix multiplication by $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

   [d]: This doesn't satisfy additivity ($(0, 1)$ and $(1, 0)$ still work). Furthermore, scaling doesn't work either (if we scale by -1 we now pick out the negative of the smallest component, which doesn't have to be related in any way to the largest component.

9. Do problem 27 from section 7.2.

Solution

The question statement is kinda confusing. I'm parsing it as: "Suppose some linear transformation T sends a basis of v_i to a basis of w_i via $T(v_i) = w_i$. Why must T be invertible?"

T is invertible because we can give an explicit inverse from its image: take the $w_i = T(v_i)$ and construct the map T' that sends w_i to v_i. This is a well-defined map because there is only one way to define what T does on any vector w (since w_i form a basis there is only one way to decompose w into w_i, which is the heart of the problem). This is easily checked to be linear.