Selected Solutions to Homwork 10

Tim Smits

8.1.5. Orthogonally diagonalize $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

Solution: We have $p_A = -(\lambda - 2)(\lambda + 1)^2$, so A has eigenvalues 2 and -1. Bases of the eigenspaces are given by $E_2 = \text{Span}\{(1,1,1)\}$ and $E_{-1} = \text{Span}\{(-1,0,1),(-1,1,0)\}$. Running Gram Schmidt gives orthonormal bases of $E_2 = \text{Span}\{(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}$ and $E_{-1} = \text{Span}\{(-1/\sqrt{2}, 0, 1/\sqrt{2}), (-1/\sqrt{6}, \sqrt{2/3}, -1/\sqrt{6})\}$. This gives $S = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & \sqrt{2/3} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}$ and $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, and $A = SDS^t$.

8.1.23. Let A be a symmetric $n \times n$ orthogonal matrix. What can you say about the eigenvalues of A? Interpret $T(x) = Ax$ geometrically when $n = 2$ and $n = 3$.

Solution: Since A is symmetric, $A = A^t$ and since A is orthogonal, $A^{-1} = A^t$, so this says $A = A^{-1}$, i.e. $A^2 = I$. If λ is an eigenvalue of A, we must therefore have $\lambda^2 = 1$ so $\lambda = \pm 1$. Since A is orthogonally diagonalizable, without loss of generality choose an orthonormal eigenbasis $\beta = \{v_1, \ldots, v_n\}$ such that $v_1, \ldots, v_k$ are associated to $\lambda = 1$ and $v_{k+1}, \ldots, v_n$ are associated to $\lambda = -1$. Then A_β is a diagonal matrix with k 1's and $n - k$ -1's, i.e. A represents a reflection around the subspace spanned by the vectors $v_{k+1}, \ldots, v_n$. Concretely for $n = 2$, A is a reflection around some line (the eigenvector associated to $\lambda = 1$), and for $n = 3$, A is a reflection around either a line (if there is only 1 eigenvector associated to 1), or a plane (if there are two).

8.2.5. Determine the definiteness of the quadratic form $q(x, y) = x^2 + 4xy + y^2$.

Solution: The matrix associate to q is given by $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, which has eigenvalues 3 and -1. This then says that q is indefinite.

8.2.9. Let A be a skew-symmetric $n \times n$ matrix, i.e. $A^t = -A$.

(a) If A is skew-symmetric, is A^2 symmetric or skew-symmetric?

(b) What can you say about the definiteness of A^2? The eigenvalues?

(c) What can you say about the complex eigenvalues of A? Which such matrices A are diagonalizable over $\mathbb{R}$?

Solution:

(a) $A^2 = -AA^t$, so $(A^2)^t = -AA^t = A^2$ says A^2 is symmetric.

(b) $x^tA^2x = -x^tAA^tx = -\|Ax\|^2 \leq 0$, so A^2 is negative semi-definite, so that all it's eigenvalues are negative.

(c) Eigenvalues of A are just the square roots of the eigenvalues of A^2. Since A^2 has non-positive eigenvalues, A has purely imaginary eigenvalues (or just 0). The only way that A can then be diagonalizable is if all eigenvalues are 0, i.e. A is the 0 matrix.

8.2.27. Consider a quadratic form $q(x) = x^tAx$, for A an $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$. Determine the image of S^{n-1} under q.

Solution: Let $\beta = \{v_1, \ldots, v_n\}$ be the corresponding orthonormal eigenbasis for A. In β-coordinates, we have $q(c_1, \ldots, c_n) = \lambda_1 c_1^2 + \ldots + \lambda_n c_n^2$. Vectors in S^{n-1} satisfy $c_1^2 + \ldots + c_n^2 = 1$, so from this constraint we see that the maximal and minimal values that q can take are λ_1 and λ_n respectively. Since q is continuous, the image of S^{n-1} under q is an interval in $\mathbb{R}$, and this interval is then therefore $[\lambda_n, \lambda_1]$.

8.3.11. Compute the SVD of $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$.

Solution: $A^tA = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 1$, so A has singular values $\sigma_1 = 2$ and $\sigma_2 = 1$, which says $\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Orthonormal bases of the eigenspaces of A^tA are given by $E_4 = \text{Span}\{(0,1)\}$ and $E_1 = \text{Span}\{(1,0)\}$, so $V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = V^t$. We have $u_1 = \frac{1}{\sigma_1}Av_1$ and $u_2 = \frac{1}{\sigma_2}Av_2$, so $u_1 = (0,1,0)$ and $u_2 = (1,0,0)$. We clearly see $u_3 = (0,0,1)$ makes $\{u_1, u_2, u_3\}$ an orthonormal basis, so we get $U = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Note: A is "almost" a Σ matrix – you just need to swap the entries 1 and 2. You can do this by swapping the columns of A, and then swapping the first two rows of A. The first operation is given by the matrix V, and the second by the matrix U. If you know how to write down permutation matrices, you could do the computation extremely quickly!

8.3.17. Let A be an $n \times m$ matrix with $\text{rank}(A) = m$ and SVD $A = U\Sigma V^t$. Show that the least squares solution to $Ax = b$ is given by $x^* = \frac{b \cdot u_1}{\sigma_1}v_1 + \ldots + \frac{b \cdot u_m}{\sigma_m}v_m$.

Solution: The least squares solution is given by solving the normal equation $A^tAx^* = A^tb$. With $A = U\Sigma V^t$, we have $A^tA = (V\Sigma^tU^t)(U\Sigma V^t)$. Since U is orthogonal, $U^tU = I$, and one can check that $\Sigma^t\Sigma = D$, the matrix of eigenvalues of A^tA. This says $A^tA = VDV^t$. So we wish to solve $VDV^tx^* = V\Sigma^tU^tb$. Since $\text{rank}(A) = m$, A^tA is invertible, so 0 is not an eigenvalue and therefore D is invertible. This gives $x^* = (VDV^t)^{-1}\Sigma^tU^tb = VD^{-1}V^tV\Sigma^tU^tb = VD^{-1}\Sigma^tU^tb$. Writing down V in terms of columns, D^{-1} as a diagonal matrix, and Σ^t, U^t in terms of rows, doing out the matrix multiplication you'll find $x^* = \frac{b \cdot u_1}{\sigma_1}v_1 + \ldots + \frac{b \cdot u_m}{\sigma_m}v_m$ as desired.

8.3.23. Let A be an $n \times m$ matrix with SVD $A = U\Sigma V^t$. Show the columns of U are an orthonormal eigenbasis for AA^t. What are the eigenvalues? How does this relate the eigenvalues of AA^t and A^tA?

Solution: We have $AA^t = (U\Sigma V^t)(V\Sigma^t U^t) = UDU^t$, because $V^tV = I$ from V being orthogonal, and $\Sigma\Sigma^t = D$ is a diagonal matrix whose first r entries are the non-zero singular values squared and 0's elsewhere. We then have $AA^tu_i = UDU^tu_i = UDe_i = U(\sigma_i^2 e_i) = \sigma_i^2 Ue_i = \sigma_i^2 u_i$, so u_i is an eigenvector of eigenvalue $\sigma_i^2 = \lambda_i$. This says the vectors u_i form an orthonormal eigenbasis of AA^t. Since λ_i are the eigenvalues of A^tA, we have that AA^t and A^tA have the same eigenvalues.

8.3.27. Let λ be a real eigenvalue of an $n \times n$ matrix A. Show that $\sigma_n \leq |\lambda| \leq \sigma_1$.

Solution: Let $\beta = \{v_1, \ldots, v_n\}$ be an orthonormal eigenbasis for A^tA,and let $\{u_1, \ldots, u_n\}$ be the corresponding basis of $\mathbb{R}^n$ whose columns form the matrix U as in the SVD of A. For any unit vector $x \in \mathbb{R}^n$, write $x = c_1v_1 + \ldots + c_nv_n$, so that $Ax = c_1Av_1 + \ldots + c_nAv_n = c_1Av_1 + \ldots + c_rAv_r$, where $v_1, \ldots, v_r$ correspond the r non-zero singular values of A. We then have $Ax = c_1\sigma_1u_1 + \ldots + c_r\sigma_ru_r$, so $\|Ax\|^2 = \sigma_1^2c_1^2 + \ldots + \sigma_r^2\lambda_r^2 = \lambda_1c_1^2 + \ldots + \lambda_rc_r^2$, where λ_i are the eigenvalues of A^tA. Since x is a unit vector, we have $c_1^2 + \ldots + c_n^2 = 1$, so $\|Ax\|^2 \leq \lambda_1(c_1^2 + \ldots + c_r^2) = \lambda_1$ and $\|Ax\|^2 \geq \lambda_n(c_1^2 + \ldots + c_r^2) = \lambda_n$ by how the eigenvalues are ordered in SVD. This then says $\sigma_n \leq \|Ax\| \leq \sigma_1$ for any unit vector $x \in \mathbb{R}^n$. If v is a unit eigenvector of A of eigenvalue λ, plugging v into the above then immediately gives $\sigma_n \leq |\lambda| \leq \sigma_1$.

8.3.31. Show that any $n \times m$ matrix A of rank r can be written as a sum of r rank 1 matrices.

Solution: Let $A = U\Sigma V^t$ be a SVD of A. By how matrix multiplication works, if $\sigma_1, \ldots, \sigma_r$ are the non-zero singular values of A, we have $A = \sigma_1u_1v_1^t + \ldots + \sigma_ru_rv_r^t$. Since $\text{im}(u_iv_i^t) \subset \text{im}(u_i)$, and $\text{rank}(u_i) = 1$ (it's just a vector!), this forces $\text{rank}(u_iv_i^t) = 1$.