Honors Linear Algebra – Problem Set 10 Solutions

May 2, 2018 in class

Problem 1

Given the dimensions of A, 2×3, the matrices U, Σ, and V in the SVD of A have dimensions 2×2, 2×3, and 3×3. Let us start by evaluating

$$A^T A = \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 8 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

The eigenvalues λ of this matrix are such that

$$\begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 8-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{vmatrix} = 0 \;\Leftrightarrow\; (1-\lambda)\left[(8-\lambda)(1-\lambda)-4\right] - 4(1-\lambda) = 0$$

This can be simplified as

$$(1-\lambda)\left[(8-\lambda)(1-\lambda)-8\right] = 0 \;\Leftrightarrow\; \lambda(1-\lambda)(\lambda-9) = 0$$

The eigenvalues of $A^T A$ are 0, 1, and 9, so the singular values of A are 1 and 3.

The eigenvector of $A^T A$ corresponding to the eigenvalue 1 is in the nullspace of

$$\begin{bmatrix} 0 & 2 & 0 \\ 2 & 7 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

Row reducing, we find

$$\begin{bmatrix} 2 & 7 & 2 \\ 0 & 2 & 0 \\ 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{7}{2} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, an eigenvector is $\mathbf{a}_1 = (-1, 0, 1)$

The eigenvector of $A^T A$ corresponding to the eigenvalue 9 is in the nullspace of

$$\begin{bmatrix} -8 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & -8 \end{bmatrix}$$

Row reducing, we find

$$\begin{bmatrix} -8 & 2 & 0 \\ 0 & -\frac{1}{2} & 2 \\ 0 & 2 & -8 \end{bmatrix} \to \begin{bmatrix} -8 & 2 & 0 \\ 0 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{4} & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, an eigenvector is $\mathbf{a}_2 = (1, 4, 1)$.

The eigenvector of A^TA corresponding to the eigenvalue 0 is in the nullspace of

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 8 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

Row reducing, we find

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, an eigenvector is $\mathbf{a}_3 = (2, -1, 2)$.

We can verify that $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are orthogonal. To make them orthonormal, we divide by their norms, and obtain the column vectors in the matrix V of the SVD:

$$\mathbf{v}_1 = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) \ , \ \mathbf{v}_2 = (\frac{1}{3\sqrt{2}}, \frac{4}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}) \ , \ \mathbf{v}_3 = (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$$

Now

$$A\mathbf{v}_1 = \mathbf{u}_1 \ \Rightarrow \ \mathbf{u}_1 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$$

$$A\mathbf{v}_2 = 3\mathbf{u}_2 \ \Rightarrow \ \mathbf{u}_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$$

We conclude that the SVD of A is

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

As we saw in class, the best rank one approximation to A is given by the rank 1 matrix of the form $\mathbf{u}_i\sigma_i\mathbf{v}_i^T$ corresponding to the largest singular value. Here, it is

$$\mathbf{u}_2\sigma_2\mathbf{v}_2^T = 3 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 2 & \frac{1}{2} \\ \frac{1}{2} & 2 & \frac{1}{2} \end{bmatrix}$$

Problem 2

V is found by computing

$$A^TA = [\mathbf{w}_1 \ \mathbf{w}_2 \ldots \mathbf{w}_n]^T[\mathbf{w}_1 \ \mathbf{w}_2 \ldots \mathbf{w}_n] = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_n^T \end{bmatrix} [\mathbf{w}_1 \ \mathbf{w}_2 \ldots \mathbf{w}_n] = \begin{bmatrix} \mathbf{w}_1^T\mathbf{w}_1 & \mathbf{w}_1^T\mathbf{w}_2 & \ldots & \mathbf{w}_1^T\mathbf{w}_n \\ \mathbf{w}_2^T\mathbf{w}_1 & \mathbf{w}_2^T\mathbf{w}_2 & \ldots & \mathbf{w}_2^T\mathbf{w}_n \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{w}_n^T\mathbf{w}_1 & \mathbf{w}_n^T\mathbf{w}_2 & \ldots & \mathbf{w}_n^T\mathbf{w}_n \end{bmatrix}$$

Now, since the columns of A are orthogonal with respective lengths $\sigma_1, \sigma_2, \ldots, \sigma_n$, $\mathbf{w}_i^T\mathbf{w}_j = 0$ unless $i = j$. When that is the case, $\mathbf{w}_i^T\mathbf{w}_i = \sigma_i^2$. Hence,

$$A^TA = \begin{bmatrix} \sigma_1^2 & 0 & \ldots & \ldots & 0 \\ 0 & \sigma_2^2 & 0 & \ldots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \sigma_n^2 \end{bmatrix}$$

This is the diagonalization of A^TA. We conclude that the eigenvector matrix of A^TA is $V = I$, the identity matrix. And the singular values of A are the square roots of the entries of the diagonal matrix above, σ_1, σ_2, ..., σ_n. Σ is the matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & \sigma_n \end{bmatrix}$$

In the SVD of A, we have

$$AV = U\Sigma \;\Leftrightarrow\; A = U\Sigma$$

where we have used that $V = I$.

Now, Σ is invertible, since its diagonal entries are nonzero, so

$$U = A\Sigma^{-1}$$

where

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & \frac{1}{\sigma_n} \end{bmatrix}$$

We conclude that the SVD of A is $A = (A\Sigma^{-1})\Sigma I$.

Problem 3

1. Consider the linear transformation T such that $T(\mathbf{u}_1) = \mathbf{u}_2$ and $T(\mathbf{u}_2) = \mathbf{u}_1$. $T \neq I$ yet we have $T(T(\mathbf{u}_1)) = T(\mathbf{u}_2) = \mathbf{u}_1$. Likewise, $T(T(\mathbf{u}_2)) = T(\mathbf{u}_1) = \mathbf{u}_2$. So $T^2 = I$ and T is its own inverse.

2. Consider the projection onto $\mathbf{u}_1$, defined by $T(\mathbf{u}_1) = \mathbf{u}_1$, and $T(\mathbf{u}_2) = \mathbf{0}$. We have $T(T(\mathbf{u}_1)) = T(\mathbf{u}_1) = \mathbf{u}_1$ and $T(T(\mathbf{u}_2)) = T(\mathbf{0}) = \mathbf{0}$ (since T is a linear transformation). Therefore, $T^2 = T$.

3. If T satisfies 2., then $T^2 = T$. If T also satisfies 1., $T^2 = I$. Combining the two, this means that $T = I$. So it is impossible to find a linear transformation such that $T \neq I$ which satisfies both 1. and 2.

Problem 4

The columns of the matrix representation are the column vectors $A\mathbf{u}_1$, $A\mathbf{u}_2$, $A\mathbf{u}_3$, $A\mathbf{u}_4$ expressed in the basis $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$. Since $\mathbf{u}_3$ and $\mathbf{u}_4$ are in $N(A)$, $A\mathbf{u}_3 = A\mathbf{u}_4 = \mathbf{0}$, the last two columns of the matrix only contain 0, whatever the output basis chosen.

Now, since $A\mathbf{u}_1 = \mathbf{w}_1$, $A\mathbf{u}_1$ can be written as $(1, 0, 0)$ in the basis $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$. Likewise, since $A\mathbf{u}_2 = \mathbf{w}_2$, $A\mathbf{u}_2$ can be written as $(0, 1, 0)$ in the basis $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$.

So the matrix representation of T in these bases is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 5

Let V be the space of continuous functions $f : [-1,1] \rightarrow \mathbb{R}$. Define $P_e : V \rightarrow V$ and $P_o : V \rightarrow V$ by

$$\forall\, x \in [-1,1]\ ,\ \forall\, f \in V\ ,$$

$$(P_e f)(x) = \frac{1}{2}\left[f(x) + f(-x)\right]$$

$$(P_o f)(x) = \frac{1}{2}\left[f(x) - f(-x)\right]$$

1. Let us take f and g in V, and $(k,h) \in \mathbb{R}^2$,

   $$(P_e(kf+hg)(x)) = \frac{1}{2}\left[(kf+hg)(x) + (kf+hg)(-x)\right] = \frac{k}{2}\left[f(x)+f(-x)\right] + \frac{h}{2}\left[g(x)+g(-x)\right] = k(P_e f)(x) + h(P_e g)(x)$$

   In particilar, $(P_e 0_V)(x) = 0$. So P_e is a linear transformation.

   Analogous reasoning for P_o would show that P_o also is a linear transformation.

   Now, let $f \in V$,

   $$\begin{aligned}(P_e^2 f)(x) = \left[P_e(P_e f)\right](x) &= \frac{1}{2}\left[(P_e f)(x) + (P_e f)(-x)\right] = \frac{1}{2}\left[\frac{1}{2}(f(x)+f(-x)) + \frac{1}{2}(f(-x)+f(x))\right] \\ &= \frac{1}{4}\left[2f(x) + 2f(-x)\right] = \frac{1}{2}\left[f(x)+f(-x)\right] = (P_e f)(x) \\ (P_o^2 f)(x) = \left[P_o(P_o f)\right](x) &= \frac{1}{2}\left[(P_o f)(x) - (P_o f)(-x)\right] = \frac{1}{2}\left[\frac{1}{2}(f(x)-f(-x)) - \frac{1}{2}(f(-x)-f(x))\right] \\ &= \frac{1}{4}\left[2f(x) - 2f(-x)\right] = \frac{1}{2}\left[f(x)-f(-x)\right] = (P_o f)(x)\end{aligned}$$

   Thus $P_e^2 = P_e$ and $P_o^2 = P_o$, so both P_e and P_o are projections.

   Finally, for any $f \in V$,

   $$\left[(I-P_e)f\right](x) = f(x) - \frac{1}{2}\left[f(x)+f(-x)\right] = \frac{1}{2}\left[2f(x) - f(x) - f(-x)\right] = \frac{1}{2}\left[f(x) - f(-x)\right] = (P_o f)(x)$$

   which proves that $P_o = I - P_e$.

2. Let f be in the kernel of P_e. Then $(P_e f)(x) = 0$ for all $x \in [-1,1]$, i.e. $[(I-P_o)f](x) = 0$, which can be written as $(P_o f)(x) = f(x)$. In other words, f is left invariant by P_o. Since P_o is a projection, this means that $f \in \mathrm{Im}P_o$. Hence $\mathrm{Ker}P_e \subset \mathrm{Im}P_o$. Conversely, if $f \in \mathrm{Im}P_o$, then $(P_o f)(x) = f(x)$, so that $(I - P_o f)(x) = 0$, i.e. $f \in \mathrm{Ker}P_e$. We conslude that $\mathrm{Im}P_o \subset \mathrm{Ker}P_e$, and thus that $\mathrm{Im}P_o = \mathrm{Ker}P_e$.

   Now, since $P_o = I - P_e$, we could use the same reasoning to show that $\mathrm{Im}P_e = \mathrm{Ker}P_o$.

   We can go beyond this result. Since $P_e + P_o = I$, $V = \mathrm{Im}P_e + \mathrm{Im}P_o$. Now, $\mathrm{Im}P_o = \mathrm{Ker}P_e$. Let $f \in \mathrm{Im}P_e \cap \mathrm{Ker}P_e$. $(P_e f)(x) = f(x)$ and $(P_e f)(x) = 0_V$, so $f(x) = 0_V$, and $\mathrm{Im}P_e \cap \mathrm{Ker}P_e = \{0_V\}$, i.e. $\mathrm{Im}P_e \cap \mathrm{Im}P_o = \{0_V\}$. We conclude that

   $$V = \mathrm{Im}P_e \bigoplus \mathrm{Im}P_o$$

   which may also be written as

   $$V = \mathrm{Ker}P_o \bigoplus \mathrm{Ker}P_e$$

3. Let $n \in \mathbb{N}^*$ be an even integer. Then

   $$P_e x^n = x^n \qquad , \qquad P_o x^n = 0$$

   Let $n \in \mathbb{N}^*$ be an odd integer. Then

   $$P_o x^n = x^n \qquad , \qquad P_e x^n = 0$$

   By linearity, we conclude that for any polynomial in P_n, P_e only keeps the even monomials in that polynomial, and P_o only keeps the odd monomials.