01:640:495 - Lecture 10

Pranav Tikkawar

March 6, 2025

1. Following are two figures from the article "A singularly Valuable Decomposition: The SVD of a Matrix" by Dan Kalman. Can you explain the idea the figures are communicating?

> **Solution:** The first figure shows orthogonal projection of a basis of $\mathbb{R}^3$ onto a plane. Then it shows the operation of multiplying the projected basis to another basis.
> The second figure shows the null space of $\mathbf{A}$ mapping to the range of $\mathbf{A}^T$ and the null space of $\mathbf{A}^T$ mapping to the range of $\mathbf{A}$.

2. Suppose $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors of $\mathbf{A}^T\mathbf{A}$ that are orthogonal $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Show their images under $\mathbf{A}$ are orthogonal - that is $\mathbf{A}\mathbf{v}_1 \cdot \mathbf{A}\mathbf{v}_2 = 0$.

> **Solution:** We can see that
>
> $$\begin{aligned}\langle Av_1, Av_2\rangle &= (Av_1)^T(Av_2) \\ &= v_1^T A^T A v_2 \\ &= v_1^T \lambda_1 v_2 \\ &= \lambda_1 v_1^T v_2 \\ &= 0\end{aligned}$$

3. Let $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 2 & -1 \end{bmatrix}$ For your convenience, you are told that $\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 6 & -5 \\ -5 & 6 \end{bmatrix}$ has eigenvalues $\lambda = 11, 1$ with eigenvectors $\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ respectively a basis for the two eigenspaces (null space (kernel) of $\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I}$). Also, $\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 2 & -3 & 3 \\ -3 & 5 & -4 \\ 3 & -4 & 5 \end{bmatrix}$ has

eigenvalues $\lambda = 1, 0, 11$ with eigenvectors $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ -1 \\ 1 \end{bmatrix}$ respectively a basis for the three eigenspaces (null space (kernel) of $\mathbf{A}\mathbf{A}^T - \lambda \mathbf{I}$).

(a) Obtain a singular value decomposition of $\mathbf{A}$.

> **Solution:** For our SVD we have
>
> $$\Sigma = \begin{bmatrix} \sqrt{11} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$
>
> $$V = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$
>
> $$U = \begin{bmatrix} \frac{2}{\sqrt{22}} & 0 & \frac{-3}{\sqrt{10}} \\ \frac{-3}{\sqrt{22}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{10}} \\ \frac{3}{\sqrt{22}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \end{bmatrix}$$

(b) Use it to write $\mathbf{A}$ as sum of rank 1 projection matrices.

> **Solution:** We can write out the SVD as
>
> $$11 \begin{bmatrix} \frac{2}{\sqrt{22}} \\ \frac{-3}{\sqrt{22}} \\ \frac{3}{\sqrt{22}} \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + 1 \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

(c) What is the rank of $\mathbf{A}$? Is $\mathbf{A}$ a 'full rank' matrix?

> **Solution:** The rank of $\mathbf{A}$ is 2. It is not a full rank matrix.

(d) Determine the spectral norm of $\mathbf{A}$. (Recall it is the maximum $\|\mathbf{AX}\|$ among $\mathbf{X}$ with $\|\mathbf{X}\| = 1$ with $\|\cdot\|$ denoting the usual Euclidean norm (length).)

> **Solution:** The spectral norm of $\mathbf{A}$ is $\sqrt{11}$. It is the value of the largest singular value.

4. Matrices $\mathbf{A}, \mathbf{b}$ are given matrices of sizes $m \times n$, $m \times 1$ respectively. And, $\mathbf{x} \in \mathbb{R}^n$ variable. Consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2$.

(a) Express $f(\mathbf{x})$ using matrix operations.

Solution:

$$
\begin{aligned}
f(x) &= (Ax-b)^T(Ax-b) \\
&= x^TA^TAx + b^Tb - b^TAx - x^TA^Tb \\
&= \langle Ax, Ax\rangle + \langle b, b\rangle - \langle b, Ax\rangle - \langle Ax, b\rangle \\
&= \langle Ax, Ax\rangle + \langle b, b\rangle - 2\langle b, Ax\rangle \\
&= ||Ax|| + ||b|| - 2\langle b, Ax\rangle
\end{aligned}
$$

This is essentially splitting up the norm into its components.

(b) Assuming the derivative exists, compute the derivative (gradient) linear map $\frac{df}{d\mathbf{x}}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$. You may use $\mathbf{h} \in \mathbb{R}^n$ as the argument of this linear map (so, for the "direction"). Examining the expression of this linear map, also determine the $1 \times n$ matrix of this linear map $\frac{df}{d\mathbf{x}}(\mathbf{x})$. *Hint: Differentiate a curve passing through* $\mathbf{x}$ *in the direction* $\mathbf{h}$ *i.e, use* $\frac{df}{d\mathbf{x}}(\mathbf{x})(\mathbf{h}) = \frac{d}{dt}\Big|_{t=0} f(\mathbf{x}+t\mathbf{h})$ *and previous part.*

Solution: Using $x+th$ as x in the previous part, we get

$$
\begin{aligned}
f(x+th) &= ||A(x+th)|| + ||b|| - 2\langle b, A(x+th)\rangle \\
&= x^TA^TAx + tx^TA^TAh + th^TA^TAx + t^2h^TA^TAh - 2b^TAx - 2tb^TAh + b^Tb
\end{aligned}
$$

Taking the derivative of this with respect to t and evaluating at $t=0$ gives us the derivative.

$$
\begin{aligned}
\frac{d}{dt}\Big|_{t=0} f(x+th) &= \frac{d}{dt}\Big|_{t=0} (||A(x+th)|| + ||b|| - 2\langle b, A(x+th)\rangle) \\
&= \frac{d}{dt}\Big|_{t=0} x^TA^TAx + tx^TA^TAh + th^TA^TAx + t^2h^TA^TAh - 2b^TAx - 2tb^TAh + b^T \\
&= x^TA^TAh + h^TA^TAx - 2th^TA^TAh - 2b^TAh\Big|_{t=0} \\
&= 2x^tA^TAh - 2b^TAh \\
&= 2(Ax-b)^TAh
\end{aligned}
$$

Thus we get

$$
\frac{d}{dx}f(x) = 2(Ax-b)^TA
$$

(c) Determine $\mathbf{x}$ where the linear map $\frac{df}{d\mathbf{x}}$ is zero. Do you recognize this scenario from earlier?

Solution: The linear map is zero then $(Ax-b)^T A = 0$. This is the normal equation for the least squares problem.

$$\begin{aligned}(Ax-b)^T A &= 0\\ A^T A x &= A^T b\\ x &= (A^T A)^{-1} A^T b\end{aligned}$$