MTH 100 : Lecture 39

Example for SVD

Let $A = \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}_{3\times 2}$

Then $A^T A = \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \end{bmatrix} \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$

$= \begin{bmatrix} 81 & -27 \\ -27 & 9 \end{bmatrix}_{2\times 2} = B$ (say) (symmetric 2x2 matrix)

Characteristic Polynomial of B

$= \det[B - \lambda I] = \begin{vmatrix} 81-\lambda & -27 \\ -27 & 9-\lambda \end{vmatrix}$

$= (81-\lambda)(9-\lambda) - (-27)(-27)$

$= \cancel{729} - 9\lambda - 81\lambda + \lambda^2 - \cancel{729}$

$= \lambda^2 - 90\lambda = \lambda(\lambda - 90)$

Eigen values in descending order are

$\lambda_1 = 90, \quad \lambda_2 = 0$

For $\lambda_1 = 90$

$$B - \lambda_1 I = \begin{bmatrix} 81-90 & -27 \\ -27 & 9-90 \end{bmatrix} = \begin{bmatrix} -9 & -27 \\ -27 & -81 \end{bmatrix}$$

$$\Big\downarrow R_2 \to R_2 - 3R_1$$

$$\text{RREF matrix} = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \xleftarrow{R_1 \to -\frac{1}{9}R_1} \begin{bmatrix} -9 & -27 \\ 0 & 0 \end{bmatrix}$$

So, the system of equation becomes:

$$\left.\begin{aligned} x_1 + 3x_2 &= 0 \\ x_2 &= x_2 \end{aligned}\right\} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

Taking $x_2 = -1$ and normalising we get an eigenvector $v_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} \end{bmatrix}$

For $\lambda_2 = 0$

$$B - \lambda_2 I = \begin{bmatrix} 81 & -27 \\ -27 & 9 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} -27 & 9 \\ 81 & -27 \end{bmatrix}$$

$$\Big\downarrow R_2 \to R_2 + 3R_1$$

$$\text{RREF matrix} = \begin{bmatrix} 1 & -\frac{1}{3} \\ 0 & 0 \end{bmatrix} \xleftarrow{R_1 \to -\frac{1}{27}R_1} \begin{bmatrix} -27 & 9 \\ 0 & 0 \end{bmatrix}$$

So, the system of equations becomes $\left.\begin{aligned} x_1 - \tfrac{1}{3}x_2 &= 0 \\ x_2 &= x_2 \end{aligned}\right\}$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} \frac{1}{3} \\ 1 \end{bmatrix}$$

Taking $x_2 = 3$ and normalising we get an eigenvector $v_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$$= \begin{bmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix}$$

Note that $\langle v_1, v_2 \rangle = 0$ (They are eigenvectors of distinct values.)

So, $V = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}_{2 \times 2}$

Note that $\sigma_1 = \sqrt{90} = 3\sqrt{10}$

$\sigma_2 = 0$

Now we will to compute U:

$$A v_1 = \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} -\frac{10}{\sqrt{10}} \\ \frac{20}{\sqrt{10}} \\ \frac{20}{\sqrt{10}} \end{bmatrix}$$

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{3\sqrt{10}} \begin{bmatrix} -\frac{10}{\sqrt{10}} \\ \frac{20}{\sqrt{10}} \\ \frac{20}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

Note that $\|u_1\| = 1$

Now $Av_2 = 0.v_2 = 0$

So, we need to extend u_1 to an orthonormal basis of $\mathbb{R}^3$ by solving the system $\langle u_1, x \rangle = u_1 . x = 0$

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be such that $\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} . x = 0$

$$\Rightarrow \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\Rightarrow -\frac{1}{3} x_1 + \frac{2}{3} x_2 + \frac{2}{3} x_3 = 0$$

$$\Rightarrow -x_1 + 2x_2 + 2x_3 = 0$$

$$\Rightarrow \left.\begin{aligned} x_1 &= 2x_2 + 2x_3 \\ x_2 &= x_2 \\ x_3 &= x_3 \end{aligned}\right\}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\left\{\begin{aligned} &\text{Taking } x_2 = 1, x_3 = 0 \text{ we get } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \\ &\text{Taking } x_2 = 0, x_3 = 1 \text{ we get } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \end{aligned}\right\}$$

Now $\left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\rangle$

$= 2 \times 2 + 0 + 0 = 4 \neq 0$

Thus any two solution may not be orthogonal to each other.
If necessary we will have to use Gram-Schmidt orthonormalisation process.

In this problem, we will find the orthonormal vectors by inspection.

Taking $x_2 = \frac{2}{3}$ and $x_3 = -\frac{1}{3}$, we get $u_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}$

Note that $\|u_2\| = 1$

Taking $x_2 = -\frac{1}{3}$ and $x_3 = \frac{2}{3}$, we get $u_3 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$

Note that $\|u_3\| = 1$

So, $$U = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}_{3\times 3}$$

$$\Sigma = \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{3\times 2}$$

With this U, V and Σ we have

$$A = U\Sigma V^T = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}_{3\times 3} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{3\times 2} \begin{bmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}_{2\times 2}$$

<u>Check:</u> $A = U\Sigma V^T$ or equivalently <u>$AV = U\Sigma$</u>:

Now, $$AV = \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}_{3\times 2} \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}_{2\times 2} = \begin{bmatrix} -\frac{10}{\sqrt{10}} & 0 \\ \frac{20}{\sqrt{10}} & 0 \\ \frac{20}{\sqrt{10}} & 0 \end{bmatrix}_{3\times 2}$$

$$= \begin{bmatrix} -\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \end{bmatrix}_{3\times 2}$$

$$U\Sigma = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}_{3\times 3} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{3\times 2}$$

$$= \begin{bmatrix} -\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \end{bmatrix}_{3\times 2}$$

So, $AV = U\Sigma$

Proof of Singular Value Decomposition (SVD)

Theorem:

Suppose that $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A^TA with corresponding eigenvalues arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Let $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, 2, \dots, n$

Suppose that A has r nonzero singular values so that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$

Then $\{Av_1, Av_2, \dots, Av_r\}$ is an orthogonal basis for col A, (Thus rank $= r$) by a previous proposition.

Normalize each Av_i to obtain an orthonormal basis $\{u_1, \dots, u_n\}$ for col A by putting

$$u_i = \frac{Av_i}{\|Av_i\|} = \frac{1}{\sigma_i} Av_i$$

i.e. $\quad Au_i = \sigma_i u_i \quad (\text{for } i = 1, 2, \dots, r) \quad \dots\dots (1)$

If $r < m$, extend $\{u_1, \dots, u_r\}$ to an orthonormal basis of $\mathbb{R}^m$ (Here we may use Gram-Schmidt Process)

Now let $U = [u_1 \; u_2 \; \dots\dots \; u_n]$ and $V = [v_1 \; v_2 \; \dots \; v_n]$

U and V are orthogonal matrices by construction.

Now
$$AV = A[v_1 \; v_2 \; .. \; v_n] = [Av_1 \; Av_2 \; \dots\dots \; Av_n]$$
$$= [\sigma_1 u_1 \;\; \sigma_2 u_2 \; \dots\dots \; \sigma_r u_r \;\; 0 \; \dots\dots \; 0]$$

[By (1)]

Now let Σ be the $m \times n$ matrix containing an $r \times r$ diagonal matrix D with the r non-zero singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and we make D into an $m \times n$ matrix Σ (same size as A) by filling out with zeros

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$m-r$ rows

$n-r$ columns

Then $$U\Sigma = [u_1 \ u_2 \ \cdots \ u_m] \begin{bmatrix} \sigma_1 & 0 \cdots 0 & \cdots & 0 \\ 0 & \sigma_2 & & \\ 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & & \vdots & \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

$$= [\sigma_1 u_1 \ \sigma_2 u_2 \ \cdots \ \sigma_r u_r \ 0 \cdots 0] = AV$$

$\Rightarrow AV = U\Sigma$

Since V is orthogonal, $\boxed{A = U \Sigma V^T}$

$\boxed{\text{QED}}$