Notes for Friday 14/11/2014 ①

Proof of Proposition 66

→ Clearly the vectors $A\bar{v}_1, \dots, A\bar{v}_r$ belong to Col A.

→ Also, for $j > r$, we have $\|A\bar{v}_j\| = \sqrt{\lambda_j} = 0$, so $A\bar{v}_j = 0$

→ Now for $i, j \leq r$, we have:

$$A\bar{v}_i \circ A\bar{v}_j = (A\bar{v}_i)^T A\bar{v}_j$$

$$= \bar{v}_i^T (A^T A)\bar{v}_j$$

$$= v_i^T (\lambda_j \bar{v}_j) = \lambda_j (\bar{v}_i^T \bar{v}_j) = 0$$

since $\{\bar{v}_1, \dots, \bar{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

Thus, $A\bar{v}_1, \dots, A\bar{v}_r$ form an orthogonal set of non-zero vectors, and so are lin. indep.

~~But now, dim (Col A) = rank A~~

(PTO)

Finally, to show that $\{A\bar{v}_1, \dots, A\bar{v}_r\}$ is a basis, we have to show that it is a spanning set for Col A.

So suppose $\bar{y} \in \text{Col } A$.

Then $\bar{y} = A\bar{x}$ for some vector $\bar{x} \in \mathbb{R}^n$.

Expressing $\bar{x} = c_1\bar{v}_1 + \dots + c_n\bar{v}_n$,

we get $\bar{y} = A\bar{x}$

$$= A(c_1\bar{v}_1 + c_2\bar{v}_2 + \dots + c_n\bar{v}_n)$$

$$= c_1 A\bar{v}_1 + c_2 A\bar{v}_2 + \dots + c_r A\bar{v}_r,$$

since remaining terms are $\bar{0}$, as noted at the start

This completes the proof, and also shows that

$$\text{rank } A = \dim(\text{Col } A)$$

$$= r$$

= no. of non-zero singular values of A

= no. of non-zero eigenvalues of $A^T A$

Solved Example for SVD

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

Sym 2×2

$\therefore \; A^T A = \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \end{bmatrix} \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix} = \begin{bmatrix} 81 & -27 \\ -27 & 9 \end{bmatrix}$

Char. poly. is $\det(A - \lambda I) = (81-\lambda)(9-\lambda) - 729$

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

$\therefore$ the eigenvalues in descending order are $\lambda_1 = 90$, $\lambda_2 = 0$.

Putting $\lambda_1 = 90$, $A - \lambda I = \begin{bmatrix} -9 & -27 \\ -27 & -81 \end{bmatrix}$

$\longrightarrow \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$.

Solving, we get $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

and by suitable choice and normalizing:-

$\bar{v}_1 = \begin{bmatrix} \frac{3}{\sqrt{10}} \\ -1/\sqrt{10} \end{bmatrix}$

Putting $\lambda_2 = 0$ and now reducing $A - \lambda I = \begin{bmatrix} 81 & -27 \\ -27 & 9 \end{bmatrix}$,

we similarly get $\bar{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ 3/\sqrt{10} \end{bmatrix}$.

NB! $\bar{v}_1$ and $\bar{v}_2$ are orthonormal as expected.

$$\therefore V = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}$$ and

the singular values are $\sigma_1 = \sqrt{90} = 3\sqrt{10}$ and $\sigma_2 = 0$.

We now need to compute U.

$$\bar{u}_1 = \frac{1}{\sigma_1} A \bar{v}_1 = \frac{1}{3\sqrt{10}} \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} \\ \frac{-1}{\sqrt{10}} \end{bmatrix}$$

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

However, $A\bar{v}_2 = \bar{0}$. So we need to extend $\bar{u}_1$ to an orthonormal basis of $\mathbb{R}^3$ by solving the system

$$\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \bar{x} = \bar{0}$$

or $x_1 = 2x_2 + 2x_3$

or $\bar{x} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} x_3$.

We need to get an orthonormal solution (if nec. we would use Gram-Schmidt process).

Here we ~~get~~ take $x_2 = \frac{2}{3},\ x_3 = -\frac{1}{3}$ ⑤

giving: $\begin{bmatrix} \frac{4}{3} - \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} = \bar{u}_2$

and $x_2 = -\frac{1}{3},\ x_3 = \frac{2}{3}$ giving

$$\begin{bmatrix} -\frac{2}{3} + \frac{4}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \bar{u}_3$$

[Check: $\bar{u}_2 \cdot \bar{u}_3 = \frac{4}{3} - \frac{2}{3} - \frac{2}{3} = 0$]

So finally: $A = U \Sigma V^T$

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

3×3 $\qquad$ 3×2 ← same as A $\qquad$ ↑ NB: Transpose of the V

Check: $AV = \begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -1/\sqrt{10} & \frac{3}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} -\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \\ 2\sqrt{10} & 0 \end{bmatrix}$

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

~~4~~ 6

A Tip for doing ~~these~~ this type of problem:

Suppose we had instead been given a 2×3 - matrix, say A_1.

Then $A_1^T A_1$ would be ~~3 3x2-matrix~~ a square 3×3 - matrix. Finding its char. poly., solving etc. would be much harder.

What is the tip?

Instead work with $B_1 = A_1^T$ → then $B_1^T B$ would be $(A_1^T)^T A_1^T$

$= A_1 A_1^T$ → i.e. $(2 \times 3) \times (3 \times 2)$ or 2×2 - matrix, easier.

Suppose we solve to get

$$B_1 = U_1 \Sigma V_1^T \quad (1)$$

Then, $A_1 = B_1^T = (U_1 \Sigma V_1^T)^T$

$$= V_1 \Sigma^T U_1^T \quad (2)$$

which we can easily get from (1), which is already solved!!

Proof of SVD Theorem (Theorem 14):-

Let λ_i and $\bar{v}_i$ be as in Proposition 66.
Thus, the $\bar{v}_i$ form an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$.
Then $\{A\bar{v}_1, \dots, A\bar{v}_r\}$ is an orthogonal basis for Col A, which is a subspace of $\mathbb{R}^m$.
Normalize each $A\bar{v}_i$ to obtain an orthonormal basis $\{\bar{u}_1, \bar{u}_2, \dots, \bar{u}_r\}$

so that $\bar{u}_i = \dfrac{A\bar{v}_i}{\|A\bar{v}_i\|} = \dfrac{1}{\sigma_i} A\bar{v}_i$

Hence $A\bar{v}_i = \sigma_i \bar{u}_i$ ①

Now, extend $\{\bar{u}_1, \dots, \bar{u}_r\}$ to an orthonormal basis of $\mathbb{R}^m$.

Put $U = [\bar{u}_1 \dots \bar{u}_m]$ and $V = [\bar{v}_1 \dots \bar{v}_n]$

[NB: U and V are orthogonal matrices by construction].

Also $AV = A[\bar{v}_1 \dots \bar{v}_n] = [A\bar{v}_1 \dots A\bar{v}_n]$

$= [\sigma_1\bar{u}_1 \;\; \sigma_2\bar{u}_2 \dots \sigma_r\bar{u}_r \;\; \bar{0} \dots \bar{0}]$ ②

(using ①)

(PTO)

Now, let D be $r \times r$ diagonal matrix with diagonal entries $\sigma_1, \sigma_2, \dots, \sigma_r$ and we make D into an $m \times n$-matrix Σ (same size as A) by filling out with zeroes.

$$\Sigma = \begin{array}{c} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \longrightarrow m-r \text{ rows} \\ \uparrow \\ n-r \text{ columns} \end{array}$$

~~Proof~~ Then: $U\Sigma = \begin{bmatrix} \bar{u}_1 \cdots \bar{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & \bar{0} \end{array}\right]$

$= \begin{bmatrix} \sigma_1 \bar{u}_1 & \cdots & \sigma_r \bar{u}_r & \bar{0} & \cdots & \bar{0} \end{bmatrix}$

$= AV$ from (2)

Since V is orthogonal:

$$U \Sigma V^T = AVV^T = A$$

as reqd.