Lecture 4:

Image decomposition

Image decomposition based on Singular Value Decomposition (SVD)

Definition: (SVD) For any $g \in \mathbb{R}^{m \times n}$, the singular value decomposition (SVD) of g is a matrix factorization: $g = U \Sigma V^T$, where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are unitary, Σ is a diagonal matrix ($\Sigma_{ij} = 0$ if $i \neq j$) with diagonal entries given by: $\underbrace{\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r}_{\text{singular values}} > 0$ with $r \leq \min(m, n)$.

Remark:

1. $g g^T = U \Sigma V^T V \Sigma^T U^T = U D U^T$ $\quad \left(D = \Sigma \Sigma^T = \text{diagonal} \right)$

   $\therefore$ U consists of eigenvectors of $g g^T$

2. Similarly, V consists of eigenvectors of $g^T g$.

Theorem: (Existence of SVD) Every $m \times n$ image has a SVD.

Proof: Consider the case when $m \leq n$.

We need the following theorem.

Theorem: Let $B \in M_{n \times n}$ be a real symmetric matrix. Then, $\exists$ orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ with corresponding eigenvalues such that

$$B = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2^T & - \\ & \vdots & \\ - & \vec{v}_n^T & - \end{pmatrix}$$

Note that $gg^T \in M_{m \times m}$ and $g^Tg \in M_{n \times n}$ are symmetric.

$\therefore \exists$ n pairwise orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ of g^Tg.

Suppose $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ are associated with non-zero eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_r$.

Note that $gg^T(g\vec{v}_i) = g(\lambda_i \vec{v}_i) = \lambda_i (g\vec{v}_i)$.

$\therefore$ $g\vec{v}_i$ is an eigenvector of gg^T with eigenvalue λ_i.

Note that g^Tg is semi positive-definite and hence all of its eigenvalues must not be -ve

$\therefore$ $\lambda_i > 0$ for $i = 1, 2, \ldots, r$.

Let $\sigma_i = \sqrt{\lambda_i}$. Then: $\|g\vec{v}_i\|^2 = (g\vec{v}_i)^T(g\vec{v}_i) = \vec{v}_i^T g^T g \vec{v}_i = \vec{v}_i^T(\lambda_i \vec{v}_i) = \lambda_i$.

$$\therefore \quad \|g\vec{v}_i\| = \sigma_i$$

Define $\vec{u}_i = \dfrac{g\vec{v}_i}{\sigma_i}$. Then: $\|\vec{u}_i\| = 1$.

$$\text{Also,} \quad \vec{u}_i \cdot \vec{u}_j = \frac{(g\vec{v}_i)^T(g\vec{v}_j)}{\sigma_i\sigma_j} = \frac{\vec{v}_i^T g^T g \vec{v}_j}{\sigma_i \sigma_j} = \frac{\lambda_j \vec{v}_i^T \vec{v}_j}{\sigma_i \sigma_j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\vec{v}_i^T \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\vec{u}_i \cdot g\vec{v}_j = \sigma_j \vec{u}_i^T \vec{u}_j = \begin{cases} \sigma_i & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

In matrix form,

$$\underbrace{\begin{pmatrix} - & \vec{u}_1^T & - \\ - & \vec{u}_2^T & - \\ & \vdots & \\ - & \vec{u}_r^T & - \end{pmatrix}}_{r \times m} \underbrace{g}_{m \times n} \underbrace{\begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_r \\ | & | & & | \end{pmatrix}}_{n \times r} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

Extend $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r\}$ to an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}_{r+1}, \dots, \vec{u}_m\}$ of $\mathbb{R}^m$.

Then:
$$\underbrace{\begin{pmatrix} - & \vec{u}_1^T & - \\ - & \vec{u}_2^T & - \\ & \vdots & \\ - & \vec{u}_r^T & - \\ & \vdots & \\ - & \vec{u}_m^T & - \end{pmatrix}}_{m \times m} \underbrace{g}_{m \times n} \underbrace{\begin{pmatrix} | & & | & & | \\ \vec{v}_1 & \cdots & \vec{v}_r & \cdots & \vec{v}_n \\ | & & | & & | \end{pmatrix}}_{n \times n} = \underbrace{\begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}}_{m \times n} \neq \Lambda^{1/2}$$

(Here, we need to use the fact that $g\vec{v}_j = 0$ for $j > r$, since $\|g\vec{v}_j\| = \sigma_j = 0$ for $j > r$)

Note that $U^T U = U U^T = I$; $V^T V = V V^T = I$. $\therefore$ $g = U \Lambda^{1/2} V^T$, where

$$\Lambda = \begin{pmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

(The case for $m > n$ can be shown similarly)

How to compute SVD

Let $A \in M_{m\times n}$

Step 1: Find eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and orthonormal eigenvectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ of $A^T A \in M_{n\times n}$ (with $\|\vec{v}_j\| = 1$, $j = 1, \dots, n$)

[Recall: $(A^T A)\vec{v}_j = \lambda_j \vec{v}_j$]

Step 2: Define:

$$\Sigma = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \\ \hline & & 0 & \end{pmatrix} \in M_{m\times n}$$

Add zero rows if $m > n$

Step 3: For non-zero $\sigma_1, \sigma_2, \dots, \sigma_r$,

let $\vec{u}_1 = \dfrac{A\vec{v}_1}{\sigma_1}$, $\vec{u}_2 = \dfrac{A\vec{v}_2}{\sigma_2}$, $\dots$, $\vec{u}_r = \dfrac{A\vec{v}_r}{\sigma_r}$

Step 4: Extend $\{\vec{u}_1, \dots, \vec{u}_r\}$ to the basis $\{\vec{u}_1, \dots, \vec{u}_r, \dots, \vec{u}_m\}$ of $\mathbb{R}^m$.

Step 5: Let:

$$U = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{pmatrix} \in M_{m\times m}$$

$$V = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{pmatrix} \in M_{n\times n}$$

Then: $A = U \Sigma V^T$

Example 2.1: Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix} .$$

We have

$$A^T A = \begin{pmatrix} 9 & 8 \\ 8 & 9 \end{pmatrix} .$$

Now, eig(A^*A) are 17 and 1, and so $\sigma_1 = \sqrt{17}$, $\sigma_2 = 1$ and

$$\Sigma = \begin{pmatrix} \sqrt{17} & 0 \\ 0 & 1 \end{pmatrix} .$$

Moreover,

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} , \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

This gives

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

$$u_i = \frac{A\vec{v}_i}{\sigma_i}$$

Since

$$\sigma_1 \vec{u}_1 = A\vec{v}_1 ,$$

we have

$$\vec{u}_1 = \frac{1}{\sqrt{17}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{34}} \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} .$$

Similarly, we have

$$\vec{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

The matrix U is, therefore, given by

$$U = \begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{-1}{\sqrt{2}} & \\ \frac{4}{\sqrt{34}} & 0 & \mathbf{u}_3 \\ \frac{3}{\sqrt{34}} & \frac{1}{\sqrt{2}} & \end{pmatrix}$$

for some vector $\mathbf{u}_3$ orthonormal to both $\mathbf{u}_1$ and $\mathbf{u}_2$. One possibility is

$$\vec{u}_3 = \frac{1}{\sqrt{17}} \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix} .$$

Finally, the SVD of A is given by

$$\begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{-1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \\ \frac{4}{\sqrt{34}} & 0 & \frac{-3}{\sqrt{17}} \\ \frac{3}{\sqrt{34}} & \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \end{pmatrix} \begin{pmatrix} \sqrt{17} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} .$$

Definition: For any k $(0 \le k \le r)$, we define

$$g_k = \sum_{j=1}^{k} \sigma_j \vec{u}_j \vec{v}_j^{\,T} \quad (\text{rank-}k \text{ approximation of } g)$$

$$\Vert$$

$$\underbrace{U \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_k & \\ & & & & \ddots \end{pmatrix} V^T}_{\text{Rank-}k\,!!}$$

Error of the approximation by SVD

Theorem: Let $f = \sum_{j=1}^{r} \sigma_j \vec{u}_j \vec{v}_j^T$ be the SVD of a $M \times N$ image f. For any $k < r$, and $f_k = \sum_{j=1}^{k} \sigma_j \vec{u}_j \vec{v}_j^T$, we have: $\| f - f_k \|_F^2 = \sum_{i=k+1}^{r} \underbrace{\sigma_i^2}_{\lambda_i}$

Proof: Let $f = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$.

Let $D \equiv f - f_k = \sum_{i=k+1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \in M_{M \times N}$.

Then, the m-th row, n-th col entry of D is given by:

$$D_{mn} = \sum_{i=k+1}^{r} \sigma_i u_{im} v_{in} \in \mathbb{R} \quad \text{where} \quad \vec{u}_i = \begin{pmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iM} \end{pmatrix} ; \quad \vec{V}_i = \begin{pmatrix} V_{i1} \\ V_{i2} \\ \vdots \\ V_{iN} \end{pmatrix}$$

$$\therefore D_{mn}^2 = \left(\sum_{i=k+1}^{r} \sigma_i u_{im} v_{in} \right)^2 = \sum_{i=k+1}^{r} \sigma_i^2 u_{im}^2 v_{in}^2 + 2 \sum_{i=k+1}^{r} \sum_{\substack{j=k+1 \\ j \neq i}}^{r} \sigma_i \sigma_j u_{im} v_{in} u_{jm} v_{jn}.$$

Thus, $\|D\|_F^2 = \sum_m \sum_n D_{mn}^2$

$$= \sum_m \sum_n \left(\sum_{i=k+1}^{r} \sigma_i^2 u_{im}^2 v_{in}^2 + 2 \sum_{i=k+1}^{r} \sum_{\substack{j=k+1 \\ j \neq i}}^{r} \sigma_i \sigma_j u_{im} v_{in} u_{jm} v_{jn} \right)$$

$$= \sum_{i=k+1}^{r} \sigma_i^2 \overbrace{\sum_m u_{im}^2}^{1} \overbrace{\sum_n v_{in}^2}^{1} + 2 \sum_{i=k+1}^{r} \sum_{\substack{j=k+1 \\ j \neq i}}^{r} \sigma_i \sigma_j \overbrace{\sum_m u_{im} u_{jm}}^{0} \overbrace{\sum_n v_{in} v_{jn}}^{0}$$

$$= \sum_{i=k+1}^{r} \underbrace{\sigma_i^2}_{=\lambda_i}$$

Remark:

- To approximate an image using SVD, arrange the eigenvalues λ_i in decreasing order and remove the last few terms in $\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$
- rank-k approximation is the optimal approximation using k-terms (in term of F-norm) (or with rank-k image)