Lecture 3

Recap: Error between images : $\|f - g\|$. Need to define a norm.

Another commonly used matrix norm

Definition: (Frobenius norm) $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$

Let $\vec{a}_j$ = j-th col of A. We have: $\|A\|_F = \sqrt{\sum_{j=1}^{n} \|a_j\|_2^2} = \sqrt{tr(A^* A)} = \sqrt{tr(AA^*)}$

where $tr(\cdot)$ = trace of the matrix.

Theorem: The Frobenius norm (F-norm) is invariant under multiplication by orthogonal matrices.
That is, for any $A \in \mathbb{R}^{m\times n}$, and orthogonal matrix $U \in \mathbb{R}^{m\times m}$, we have:

$$\|UA\|_F = \|A\|_F.$$

Proof:

$$\|UA\|_F = \sqrt{tr((UA)^T(UA))} = \sqrt{tr(A^TU^TUA)} = \sqrt{tr(A^TA)} = \|A\|_F$$

Recap:

Image decomposition

Suppose $h(x,\alpha,y,\beta) = h_c(x,\alpha)\, h_r(y,\beta)$ (Separable).

Then: $g = h_c^T f h_r \Rightarrow f = (h_c^T)^{-1} g (h_r)^{-1}$

Write: $(h_c^T)^{-1} = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_N \\ | & | & & | \end{pmatrix}$; $h_r^{-1} = \begin{pmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2^T & - \\ & \vdots & \\ - & \vec{v}_N^T & - \end{pmatrix}$

Then: $f = \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij}\, \vec{u}_i \vec{v}_j^T$ $M_{N\times N}$

Check that: $(h_c^T)^{-1} \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} h_r^{-1} = \vec{u}_i \vec{v}_j^T$

(i,j)-entry

$\therefore f =$ linear combination of $\{\vec{u}_i \vec{v}_j^T\}_{i,j}$

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: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ with $r \leq \min(m, n)$.

Theorem: The rank of g is given by the number of non-zero singular values.

Proof: Rank = dim of column space.

Recall that $\text{rank}(AB) = \text{rank}(B)$ if A is invertible

$\text{rank}(AB) = \text{rank}(A)$ if B is invertible.

Suppose $g = U\Sigma V^T$. Since U and V are invertible, $\text{rank}(g) = \text{rank}(\Sigma)$

$=$ # of non-zero singular values

Theorem: $\mathrm{Range}(g) = \mathrm{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$ where $U = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{pmatrix}$; $r = \mathrm{rank}(g)$

$\mathrm{Null}(g) = \mathrm{span}\{\vec{v}_{r+1}, \ldots, \vec{v}_n\}$ where $V = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{pmatrix}$.

Proof: Exercise.

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

Proof: Later!

How to compute SVD

Let $A \in M_{m\times n}$ $(m > n)$

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 $A^TA \in M_{n\times n}$ (with $\|\vec{v}_j\| = 1$, $j = 1, \dots, n$)

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

Step 2: Define:

$$\Sigma = \begin{pmatrix} \overset{=\sigma_1}{\sqrt{\lambda_1}} & & & \\ & \overset{=\sigma_2}{\sqrt{\lambda_2}} & & \\ & & \ddots & \\ & & & \overset{=\sigma_n}{\sqrt{\lambda_n}} \\ \hline & & 0 & \end{pmatrix} \in M_{m\times n}$$

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^TA) 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

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

Remark:

1. Note that $gg^T = U\Lambda^{1/2}\underbrace{V^TV}_{I}\Lambda^{1/2}U^T = U\Lambda U^T$

   $\therefore$ U consists of eigenvectors of gg^T.

   Note that $g^Tg = V\Lambda^{1/2}\underbrace{U^TU}_{I}\Lambda^{1/2}V^T = V\Lambda V^T$

   $\therefore$ V consists of eigenvectors of g^Tg.

2. Note that
$$g = U\underline{\underline{\Lambda}}^{1/2}V^T = \sum_{i=1}^{r}\sigma_i\, U \begin{array}{cc} \text{ith} & \\ \downarrow & \\ \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{pmatrix} & \leftarrow \text{ith} \end{array} V^T = \sum_{i=1}^{r}\sigma_i \vec{u}_i \vec{v}_i^{\,T}$$

   $\vec{u}_i\vec{v}_i^{\,T}$ is called the eigen-image of g under SVD.

3. For $N \times N$ image, the required storage is:
$$(\underbrace{N}_{\vec{u}_i} + \underbrace{N}_{\vec{v}_i} + \underbrace{1}_{\sigma_i}) \times \underbrace{r}_{r\text{ terms}} = (2N+1)r$$

Definition: For any k $(0 \le k \le r)$, we define
$$g_k = \sum_{j=1}^{k}\sigma_j \vec{u}_j \vec{v}_j^{\,T} \quad (\text{rank-}k \text{ approximation of } g)$$

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: $\| f - f_k \|_F^2 = \sum_{i=k+1}^{r} \underbrace{\sigma_i^2}_{}$ λ_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, $\|D\|_F^2 = \sum_m \sum_n D_{mn}^2$

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

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

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

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)

Next time: Let $g \in M_{m \times n}$

① There exists orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of $g^T g \Rightarrow$

$$g^T g \, \vec{v}_i = \lambda_i \vec{v}_i$$

Assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

② Form : $\vec{u}_1 = \dfrac{g \vec{v}_1}{\sqrt{\lambda_1}}, \vec{u}_2 = \dfrac{g \vec{v}_2}{\sqrt{\lambda_2}}, \ldots, \vec{u}_r = \dfrac{g \vec{v}_r}{\sqrt{\lambda_r}}$

(i) $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$ are orthonormal

(ii) $\vec{u}_i \cdot g \vec{v}_j = \begin{cases} \sqrt{\lambda_i} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

In matrix form :
$$\begin{pmatrix} \text{---} \; \vec{u}_1^T \; \text{---} \\ \text{---} \; \vec{u}_2^T \; \text{---} \\ \vdots \\ \text{---} \; \vec{u}_r^T \; \text{---} \end{pmatrix} g \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_r \\ | & | & & | \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_r} \end{pmatrix}$$