

Lecture 3

Recall:

Convolution

Definition: Consider $k \in M_{N \times N}(\mathbb{R})$ and $f \in M_{N \times N}(\mathbb{R})$. Assume k and f are periodically extended. That is:

$$k(x, y) = k(x + pN, y + qN)$$

$$f(x, y) = f(x + pN, y + qN)$$

where p, q are integers.

The convolution $k * f$ of k and f is a $N \times N$ matrix defined

as:

$$k * f(\alpha, \beta) = \sum_{x=1}^N \sum_{y=1}^N k(x, y) f(\alpha - x, \beta - y) \quad \text{for } 1 \leq \alpha, \beta \leq N$$

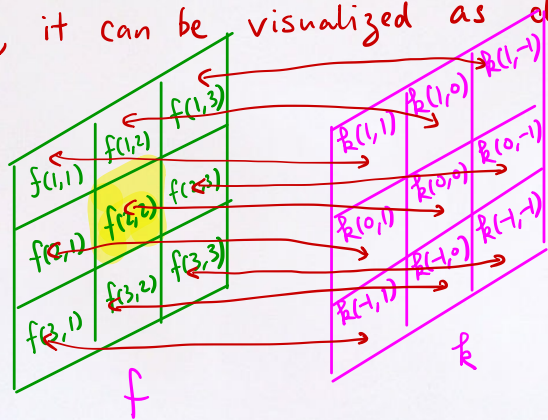
Geometric meaning of convolution

Consider $k \in M_{3 \times 3}(\mathbb{R})$ and $f \in M_{3 \times 3}(\mathbb{R})$.

$$\text{Consider: } k * f(2,2) = \sum_{x=1}^3 \sum_{y=1}^3 k(2-x, 2-y) f(x,y)$$

$$= k(1,1)f(1,1) + k(1,0)f(1,2) + k(1,-1)f(1,3) + k(0,1)f(2,1) + k(0,0)f(2,2) \\ + k(0,-1)f(2,3) + k(-1,1)f(3,1) + k(-1,0)f(3,2) + k(-1,-1)f(3,3)$$

Geometrically, it can be visualized as dot product:



Overlay k onto f
and take dot product.

Definition: The point spread function $h^{\alpha, \beta}(x, y)$ of a linear image transformation is called **shift-invariant** if there exists a function $\tilde{h}$

Such that

$$h^{\alpha, \beta}(x, y) = \tilde{h}(\alpha - x, \beta - y)$$

for all $1 \leq x, y, \alpha, \beta \leq N$.

Remark: Given $k \in M_{N \times N}(\mathbb{R})$. Let $\mathcal{O}$ be a linear image transformation defined by: $\mathcal{O}(f) = k * f$ for all $f \in M_{N \times N}(\mathbb{R})$.

Then: the point spread function of $\mathcal{O}$ is shift-invariant.

Let $g = \mathcal{O}(f)$

$$g(\alpha, \beta) = \mathcal{O}(f)(\alpha, \beta) = \sum_{x=1}^N \sum_{y=1}^N \underset{h^{\alpha, \beta}(x, y)}{k(\alpha - x, \beta - y)} f(x, y)$$

Similarity between images

Need to define matrix norm $\|\cdot\|$ such that: for $\forall f, g \in \mathcal{I}$, we can define similarity between f and g as $\|f - g\|$.

Definition: A vector/matrix norm is a function $\|\cdot\|: \mathbb{R}^m$ (or $\mathbb{R}^{m \times n}$) $\rightarrow \mathbb{R}$ so that for any $\vec{x}, \vec{y} \in \mathbb{R}^m$ (or $\mathbb{R}^{m \times n}$) and $\alpha \in \mathbb{R}$, we have:

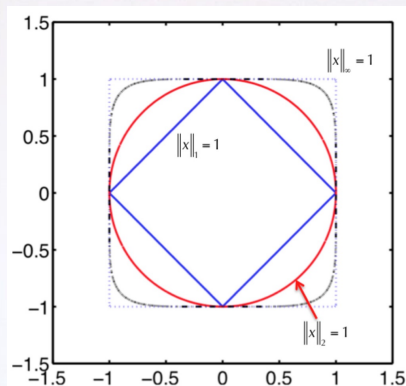
1. $\|\vec{x}\| \geq 0$, $\|\vec{x}\| = 0$ iff $\vec{x} = 0$.
2. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (triangle inequality)
3. $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$

Example:

- $\|\vec{x}\|_1 = \sum_{i=1}^m |x_i|$
- $\|\vec{x}\|_2 = \left(\sum_{i=1}^m x_i^2 \right)^{1/2}$
- $\|\vec{x}\|_\infty = \max_{i=1,2,\dots,m} |x_i|$

$\vec{x} = (x_1, x_2, \dots, x_m)^T$

} Vector norm



Remark: In image processing, vector norm can be considered as matrix norm.

Suppose $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \dots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix} \in M_{N \times N}(\mathbb{R}) \rightarrow \vec{A} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \\ \vdots \\ a_{1N} \\ a_{2N} \\ \vdots \\ a_{NN} \end{pmatrix}$

1st col of A (pointing to the first column of $\vec{A}$)
Nth col of A (pointing to the Nth column of $\vec{A}$)

Then: $\|A\|_1 = \|\vec{A}\|_1 = \sum_{i=1}^N \sum_{j=1}^N |a_{ij}|$

Given two images A and B, similarity between them can be measured by: $\sum_{i=1}^N \sum_{j=1}^N |a_{ij} - b_{ij}|$

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: $\|A\|_F = \sqrt{\sum_{j=1}^n \|\vec{a}_j\|_2^2} = \sqrt{\text{tr}(A^T A)} = \sqrt{\text{tr}(A A^T)}$
where $\text{tr}(\cdot)$ = trace of the matrix.

Importance of defining correct norm

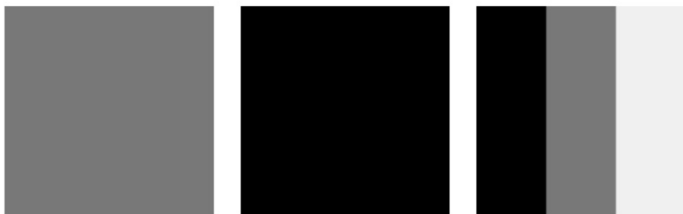


Figure 1: The images on the left and on the right are equally similar to the image in the middle in terms of the entrywise 1-norm. On the other hand, the image on the right is significant less similar to the image in the middle in terms of the entrywise 2-norm than the image on the left.

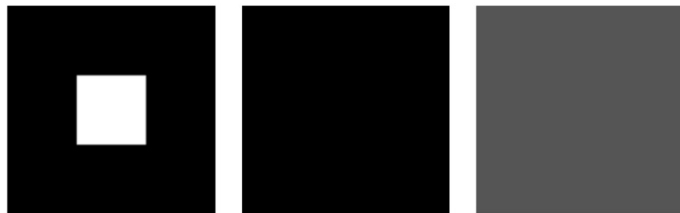


Figure 2: The images on the left and on the right are equally similar to the image in the middle in terms of the entrywise 2-norm. On the other hand, the image on the right is significant less similar to the image in the middle in terms of the entrywise 1-norm than the image on the left.

Representation of $\mathcal{O}$ by a matrix H : Let $g = \mathcal{O}(f) \in M_{N \times N}(\mathbb{R})$.

Then: $g(\alpha, \beta) = \sum_{x=1}^N \sum_{y=1}^N h^{\alpha, \beta}(x, y) f(x, y) \quad \text{for } 1 \leq \alpha, \beta \leq N$

So,

$$\begin{aligned}
 g(1,1) &= h^{1,1}(1,1)f(1,1) + \dots + h^{1,1}(N,1)f(N,1) + \dots + h^{1,1}(1,N)f(1,N) + \dots + h^{1,1}(N,N)f(N,N) \\
 g(2,1) &= h^{2,1}(1,1)f(1,1) + \dots + h^{2,1}(N,1)f(N,1) + \dots + h^{2,1}(1,N)f(1,N) + \dots + h^{2,1}(N,N)f(N,N) \\
 &\vdots \\
 g(\alpha, \beta) &= h^{\alpha, \beta}(1,1)f(1,1) + \dots + h^{\alpha, \beta}(N,1)f(N,1) + \dots + h^{\alpha, \beta}(1,N)f(1,N) + \dots + h^{\alpha, \beta}(N,N)f(N,N) \\
 &\vdots \\
 g(N,N) &= h^{N,N}(1,1)f(1,1) + \dots + h^{N,N}(N,1)f(N,1) + \dots + h^{N,N}(1,N)f(1,N) + \dots + h^{N,N}(N,N)f(N,N)
 \end{aligned}$$

(*)

N^2 equations, N^2 variables. LHS = $\begin{pmatrix} g(1,1) \\ g(2,1) \\ \vdots \\ g(N,1) \\ \vdots \\ g(1,N) \\ \vdots \\ g(N,N) \end{pmatrix} = \underset{\mathbb{R}^{N^2}}{\text{vec}} \left| \begin{array}{l} \text{Variables on RHS:} \\ \begin{pmatrix} f(1,1) \\ f(2,1) \\ \vdots \\ f(N,1) \\ \vdots \\ f(1,N) \\ \vdots \\ f(N,N) \end{pmatrix} = \underline{f} \in \mathbb{R}^{N^2} \end{array} \right.$

So, (x) can be written in matrix form:

$$\vec{g} = H \vec{f} \quad (H \in M_{N^2 \times N^2}(\mathbb{R}))$$

H is called the transformation matrix representing $\mathcal{O}$.

Example 1.1 A linear operator is such that it replaces the value of each pixel by the average of its four nearest neighbours. Assume the image is repeated in all directions. Apply this operator $\mathcal{O}$ to a 3×3 image. Find the transformation matrix corresponding to $\mathcal{O}$.

Solution:

$$\begin{array}{ccccccc}
 & & & f_{11} & f_{12} & f_{13} & \leftarrow \text{Row -3} \\
 & & & f_{21} & f_{22} & f_{23} & \leftarrow \text{Row -2} \\
 & & & f_{31} & f_{32} & f_{33} & \leftarrow \text{Row -1} \\
 3 \times 3 \text{ image} = & f_{11} & f_{12} & f_{13} & \left(\begin{array}{ccc} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{array} \right) & f_{11} & f_{12} & f_{13} & \leftarrow \text{Row 1} \\
 & f_{21} & f_{22} & f_{23} & & f_{21} & f_{22} & f_{23} & \leftarrow \text{Row 2} \\
 & f_{31} & f_{32} & f_{33} & & f_{31} & f_{32} & f_{33} & \leftarrow \text{Row 3} \\
 & \uparrow & \uparrow & \uparrow & f_{11} & f_{12} & f_{13} & \uparrow & \uparrow & \uparrow & \leftarrow \text{Row 4} \\
 & \text{Col-3} & \text{Col-2} & \text{Col-1} & f_{21} & f_{22} & f_{23} & \text{Col 4} & \text{Col 5} & \text{Col 6} & \leftarrow \text{Row 5} \\
 & & & & f_{31} & f_{32} & f_{33} & & & & \leftarrow \text{Row 6}
 \end{array}$$

$$g_{22} = \frac{f_{12} + f_{21} + f_{23} + f_{32}}{4} ; \quad g_{33} = \frac{f_{23} + f_{32} + f_{31} + f_{13}}{4}$$

etc ...

Write

$$g = \begin{pmatrix} g_{11} \\ g_{21} \\ g_{31} \\ g_{12} \\ g_{22} \\ g_{32} \\ g_{13} \\ g_{23} \\ g_{33} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{pmatrix}$$

$$g_{11} = \frac{f_{13} + f_{31} + f_{12} + f_{21}}{4}$$

$$g_{21} = \frac{f_{11} + f_{22} + f_{31} + f_{23}}{4}$$

$h^{1,1}(3,1)$ $h^{2,1}(2,3)$
etc

By careful examination, we see that

$$\begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 0 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \\ 0 & 1/4 & 0 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1/4 & 0 & 0 & 1/4 & 1/4 & 1/4 & 0 \end{bmatrix}$$

All entries are given by the point spread function
 $h^{\alpha, \beta}(x, y)$

Example Consider $\mathcal{O} : M_{N \times N}(\mathbb{R}) \rightarrow M_{N \times N}(\mathbb{R})$ defined by:

$$\mathcal{O}(f) = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} f \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ for all } f \in M_{N \times N}(\mathbb{R}).$$

Let $g = \mathcal{O}(f)$. Then:

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} f_{11} + 3f_{12} & 2f_{11} + 4f_{12} \\ f_{11} + 2f_{21} + 3f_{12} + 6f_{22} & 2f_{11} + 4f_{21} + 4f_{12} + 8f_{22} \end{pmatrix}$$

$$\begin{pmatrix} g_{11} \\ g_{21} \\ g_{12} \\ g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 2 & 3 & 6 \\ 2 & 0 & 4 & 0 \\ 2 & 4 & 4 & 8 \end{pmatrix} \begin{pmatrix} f_{11} \\ f_{21} \\ f_{12} \\ f_{22} \end{pmatrix}$$

$$\begin{pmatrix} 1 \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} & 3 \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \\ 2 \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} & 4 \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1A & 3A \\ 2A & 4A \end{pmatrix}$$

Remark: • Separable image transformation has a special structure.

- Let A and B be two matrices.

Kronecker product of A and $B = \underset{\substack{\text{"} \\ (a_{ij})_{1 \leq i, j \leq N}}}{A} \otimes B := \begin{pmatrix} a_{11} B & \dots & a_{1N} B \\ a_{21} B & \dots & a_{2N} B \\ \vdots & \vdots & \vdots \\ a_{N1} B & \dots & a_{NN} B \end{pmatrix}$

- In general, if $\mathcal{Q}: M_{N \times N}(\mathbb{R}) \rightarrow M_{N \times N}(\mathbb{R})$ is defined by:

$$\mathcal{Q}(f) = A f B \quad \text{for all } f \in M_{N \times N}(\mathbb{R}), \text{ where } A, B \in M_{N \times N}(\mathbb{R})$$

Then, the transformation matrix of $\mathcal{Q}$ is:

$$H = B^T \otimes A$$

- So, instead of storing $N^2 \times N^2 = N^4$ entries, we only need to store entries of A and B , which is $2N^2$ (much less storage)

Image decomposition

Let $g = A f B$ (Separable). Then: $f = A^{-1} g B^{-1}$

Write: $A^{-1} = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_N \\ | & | & & | \end{pmatrix}$; $B^{-1} = \begin{pmatrix} - & \vec{z}_1^T & - \\ - & \vec{z}_2^T & - \\ - & \vdots & - \\ - & \vec{z}_N^T & - \end{pmatrix}$ (assume that A and B are invertible)

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

Proof: $f = A^{-1} g B^{-1} = A^{-1} \left(\sum_{i=1}^N \sum_{j=1}^N g_{ij} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & 1 & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \leftarrow i^{\text{th}} \text{ row} \right) B^{-1}$

$= \sum_{i=1}^N \sum_{j=1}^N g_{ij} \left(A^{-1} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & 1 & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} B^{-1} \right) = \sum_{i=1}^N \vec{u}_i \vec{v}_j^T$

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

$\vec{v}_j^T$ is $j^{\text{th}} \text{ col}$
 $\vec{u}_i$ is $i^{\text{th}} \text{ row}$
 $\vec{v}_j^T$ is $N \times N \text{ matrix}$

Definition: Each $\vec{u}_i \vec{v}_j^T$ is called an elementary image.

$\vec{u}_i \vec{v}_j^T$ is also called the outer product of $\vec{u}_i$ and $\vec{v}_j$.

One important task in image processing:

Choose A and B such that:

1. Transformed image requires less storage (Many $g_{ij} = 0$)
2. Take away some terms $g_{ij} \vec{u}_i \vec{v}_j^T$ (e.g. high-frequency) $\rightarrow$ Better image!!
3. A^{-1} and B^{-1} are easy to compute!

Common example:

Orthogonal $U \Leftrightarrow U^T U = I \quad \therefore U^{-1} = U^T.$

Example: Let $A = \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}}_{A^{-1}} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}_g \underbrace{\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}}_{B^{-1}}$

Then: $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \left[1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$

$$= 1 \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} + 2 \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

$$= \underbrace{1}_{g_{11}} \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}}_{\parallel} \underbrace{\begin{pmatrix} 3 \\ 0 \end{pmatrix}}_{\parallel} + \underbrace{2}_{g_{22}} \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}}_{\parallel} \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\parallel}$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \end{pmatrix} \underbrace{\vec{v}_1^T}_{\vec{u}_1}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \end{pmatrix} \underbrace{\vec{v}_2^T}_{\vec{u}_2}$$

$$= 1 \underbrace{\begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix}}_{\text{elementary image}} + 2 \underbrace{\begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix}}_{\text{elementary image}}$$

elementary image

elementary image

Image decomposition

Image decomposition based on Singular Value Decomposition (SVD)

Definition: (SVD) For any $g \in M_{m \times n}$, the singular value decomposition (SVD) of g is a matrix factorization: $g = U \Sigma V^T$, where $U \in M_{m \times m}$, $V \in M_{n \times n}$ are orthogonal, $\Sigma \in M_{m \times n}$ 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 $r \leq \min(m, n)$. ($U U^T = U^T U = I$; $V V^T = V^T V = I$)

Singular values

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

Remark: Consider an image g . Let $g = U \Sigma V^T$ be the SVD of g (with diagonal entries of Σ given by $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$)

1. Note that $g = U \Sigma V^T = \sum_{i=1}^r \sigma_i U \begin{pmatrix} 0 & & \\ & \ddots & \\ & & 1 & \\ & & & \ddots & 0 \end{pmatrix}_{i\text{th}} V^T = \sum_{i=1}^r \sigma_i \vec{u}_i \vec{v}_i^T$

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

2. For $N \times N$ image, the required storage is:

$$(\underbrace{N}_{\vec{u}_i} + \underbrace{N}_{\vec{v}_i} + \underbrace{1}_{\sigma_i}) \times \underbrace{r}_{\text{terms}} = (2N+1)r$$

Observation about SVD Let $A = U \Sigma V^T$ (Let $A \in M_{n \times n}$)

Write $U = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_n \\ | & | & & | \end{pmatrix}$; $V = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{pmatrix}$; $\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix}$

$$\bullet A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^T \underbrace{U^T U}_I \Sigma V^T = V \Sigma^T \Sigma V^T$$

$$\Rightarrow (A^T A) V = V \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix} \Rightarrow \begin{pmatrix} | & & | \\ (A^T A) \vec{v}_1 & \dots & (A^T A) \vec{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \sigma_1^2 \vec{v}_1 & \dots & \sigma_n^2 \vec{v}_n \\ | & & | \end{pmatrix}$$

$$\Rightarrow (A^T A) \vec{v}_1 = \sigma_1^2 \vec{v}_1, \dots, (A^T A) \vec{v}_n = \sigma_n^2 \vec{v}_n$$

$\therefore \vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are eigenvectors of $A^T A$ with eigenvalues $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$.

$$\begin{aligned} \bullet \quad AA^T &= (U \Sigma V^T)(U \Sigma V^T)^T = U \Sigma \underbrace{V^T V}_I \Sigma^T U^T = U \Sigma \Sigma^T U^T \\ \Rightarrow (AA^T)U &= U \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} \Rightarrow \begin{pmatrix} AA^T \vec{u}_1 & \dots & AA^T \vec{u}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} \sigma_1^2 \vec{u}_1 & \dots & \sigma_n^2 \vec{u}_n \\ | & & | \end{pmatrix} \end{aligned}$$

$\therefore \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ are eigenvectors of AA^T with eigenvalues $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$

$$\bullet \quad A = U \Sigma V^T \Rightarrow AV = U \Sigma \Rightarrow \begin{pmatrix} A \vec{v}_1 & A \vec{v}_2 & \dots & A \vec{v}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} \sigma_1 \vec{u}_1 & \sigma_2 \vec{u}_2 & \dots & \sigma_n \vec{u}_n \\ | & | & & | \end{pmatrix}$$

$\therefore$ For $\sigma_1, \sigma_2, \dots, \sigma_r > 0$,

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

We can obtain $\vec{u}_1, \dots, \vec{u}_r$ from $\vec{v}_1, \dots, \vec{v}_r$.

$$\bullet \quad U^T U = I \Rightarrow \vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$$

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

$\therefore \{ \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \}$ are orthonormal.

$\{ \vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \}$ are orthonormal.

How to compute SVD

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

$$\begin{array}{ccc} \sigma_1^2 & \sigma_2^2 & \sigma_n^2 \\ \text{"} & \text{"} & \text{"} \end{array}$$

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

Step 3: For non-zero $\sigma_1, \sigma_2, \dots, \sigma_r$,
let $\vec{u}_1 = \frac{A \vec{v}_1}{\sigma_1}, \vec{u}_2 = \frac{A \vec{v}_2}{\sigma_2}, \dots, \vec{u}_r = \frac{A \vec{v}_r}{\sigma_r}$

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

Step 5: Let:

$$U = \begin{pmatrix} \frac{1}{\sigma_1} & \frac{1}{\sigma_2} & \dots & \frac{1}{\sigma_m} \\ \frac{1}{\sigma_1} & \frac{1}{\sigma_2} & \dots & \frac{1}{\sigma_m} \end{pmatrix} \in M_{m \times m}$$

$$V = \begin{pmatrix} \frac{1}{\sigma_1} & \frac{1}{\sigma_2} & \dots & \frac{1}{\sigma_n} \\ \frac{1}{\sigma_1} & \frac{1}{\sigma_2} & \dots & \frac{1}{\sigma_n} \end{pmatrix} \in M_{n \times n}$$

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

Example: (2x2 example) Find the SVD of $A = \begin{pmatrix} 4 & 0 \\ 3 & -5 \end{pmatrix}$

Step 1: $A^T A = \begin{pmatrix} 25 & -15 \\ -15 & 25 \end{pmatrix}$

Characteristic polynomial: $\det(A^T A - \lambda I) = (25 - \lambda)(25 - \lambda) - 15^2$
 $= \lambda^2 - 50\lambda + 400$

$\therefore A^T A$ has two eigenvalues: $\lambda = 10$ and $\lambda = 40$
 $= (\lambda - 10)(\lambda - 40)$

For $\lambda = 40$, $\vec{0} = (A^T A - 40 I) = \begin{pmatrix} -15 & -15 \\ -15 & -15 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. $\vec{v} = x_1 \begin{pmatrix} -1 \\ +1 \end{pmatrix}$

Choose $\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} / \sqrt{2} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

For $\lambda = 10$, $(A^T A - 10 I) = \begin{pmatrix} 15 & -15 \\ -15 & 15 \end{pmatrix}$

Find null space to find eigenvector.

RREF of $\begin{pmatrix} 15 & -15 \\ -15 & 15 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} = 0$

Choose $\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} / \sqrt{1^2 + 1^2} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

Let $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \text{eigenvector}$

Then: $x_1 - x_2 = 0$

$\therefore \vec{v} = x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$\therefore V = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{pmatrix}$$

Step 2: $\vec{u}_1 = \frac{A\vec{v}_1}{\sigma_1} = \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$ and $\vec{u}_2 = \frac{A\vec{v}_2}{\sigma_2} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{pmatrix}$.

$\therefore$ SVD of A is:

$$\begin{pmatrix} 4 & 0 \\ 3 & -5 \end{pmatrix} = \underbrace{\begin{pmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix}}_U \underbrace{\begin{pmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}}_{V^T}$$