

Question 3

- How will you solve for the minimum of the following objective functions: (1) $J(\mathbf{A}_r) = \|\mathbf{A} - \mathbf{A}_r\|_F^2$, where \mathbf{A} is a known $m \times n$ matrix of rank greater than r , and \mathbf{A}_r is a rank- r matrix, where $r < m, r < n$. (2) $J(\mathbf{R}) = \|\mathbf{A} - \mathbf{R}\mathbf{B}\|_F^2$, where $\mathbf{A} \in \mathbb{R}^{n \times m}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{R} \in \mathbb{R}^{n \times n}$, $m > n$ and \mathbf{R} is constrained to be orthonormal. Note that \mathbf{A} and \mathbf{B} are both known.

In both cases, explain briefly any one situation in image processing where the solution to such an optimization problem is required.

Answer:

1. We have to minimize the objective function $J(\mathbf{A}_r)$ given by,

$$J(\mathbf{A}_r) = \|\mathbf{A} - \mathbf{A}_r\|_F^2$$

such that $\text{rank}(\mathbf{A}_r) = r$.

In this problem we are basically trying to find the best rank r approximation for the matrix A . The solution to this problem is given by Eckart-Young theorem.

Suppose the singular value decomposition of the matrix A is given as,

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

then, considering the r largest singular values and the corresponding singular vectors from U and V , we can find the rank r approximation of A as,

$$\mathbf{A}_r = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^T$$

where $\mathbf{\Sigma}_r$ contains the largest r singular values of A and \mathbf{U}_r and \mathbf{V}_r are the corresponding r columns of U and V .

Application: The low rank approximation of a matrix can be used for robust patch based denoising of images/videos. Small patches from different spatial regions of an image are quite similar to each other. This similarity can be exploited to reduce the noise in the image without strong assumptions on the statistical properties of noise. Following is the link to the paper, which uses the low rank approximation to denoise an image.

Low Rank Approximation for Image Denoising

2. In this part we have to minimize the objective function $J(\mathbf{R})$ given by,

$$J(\mathbf{R}) = \|\mathbf{A} - \mathbf{R}\mathbf{B}\|_F^2$$

such that \mathbf{R} is an orthonormal matrix, i.e. $\mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{I}$.

This is the Orthogonal Procrustes problem.

Consider,

$$\begin{aligned}\|\mathbf{A} - \mathbf{RB}\|_F^2 &= \text{trace}((\mathbf{A} - \mathbf{RB})^T(\mathbf{A} - \mathbf{RB})) \\ &= \text{trace}(\mathbf{A}^T \mathbf{A} - 2\mathbf{A}^T \mathbf{RB} + \mathbf{BB}^T)\end{aligned}$$

Hence,

$$\begin{aligned}\min_{\mathbf{R}, \mathbf{B}} \|\mathbf{A} - \mathbf{RB}\|_F^2 &= \min_{\mathbf{R}, \mathbf{B}} \text{trace}(\mathbf{A}^T \mathbf{A} - 2\mathbf{A}^T \mathbf{RB} + \mathbf{BB}^T) \\ &= \max_{\mathbf{R}, \mathbf{B}} \text{trace}(\mathbf{A}^T \mathbf{RB}) \\ &= \max_{\mathbf{R}, \mathbf{B}} \text{trace}(\mathbf{RBA}^T) \dots (\because \text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})) \\ &= \max \text{trace}(\mathbf{RUSV}^T)\end{aligned}$$

This last equality is using the SVD of $\mathbf{BA}^T = \mathbf{USV}^T$

Hence,

$$\begin{aligned}\min_{\mathbf{R}, \mathbf{B}} \|\mathbf{A} - \mathbf{RB}\|_F^2 &= \max \text{trace}(\mathbf{SV}^T \mathbf{RU}) \dots (\because \text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})) \\ &= \max \text{trace}(\mathbf{SX})\end{aligned}$$

where $\mathbf{X} = \mathbf{V}^T \mathbf{RU}$ is an orthogonal matrix.

Now,

$$\text{trace}(\mathbf{SX}) = \sum_i S_{ii} X_{ii}$$

Since, values S_{ii} are non-negative, and the above expression is maximized if $X_{ii} = 1$ all along its diagonal. As \mathbf{X} is orthonormal, we must have

$$\begin{aligned}\mathbf{X} &= \mathbf{I} \\ \implies \mathbf{V}^T \mathbf{RU} &= \mathbf{I} \\ \implies \mathbf{R} &= \mathbf{VU}^T\end{aligned}$$

Application : We saw one application of this problem in Tomography. In tomography under unknown angles for 3D images, we are trying to find the best rotation matrix \mathbf{R} such that the error in the reconstruction is minimized. In determining the best rotation matrix, we use the Orthogonal Procrustes problem.