

 2 Show that $\bar{X} = X(I - \frac{1}{n}ee^T)$ (here e = vector of all ones).
What does the projector $(I - \frac{1}{n}ee^T)$ do?

Solution: Each column of \bar{X} is $\bar{x} = x - \mu$ so that $\bar{X} = X - \mu e^T$, where μ is the sample mean. But we have $\mu = \frac{1}{n} \sum x_i = \frac{1}{n} X e$ and so,

$$\bar{X} = X - \frac{1}{n} X e e^T = X [I - \frac{1}{n} e e^T]$$

The matrix $(I - \frac{1}{n}ee^T)$ represents a projector that centers the data so the mean is zero.

 3 Show that solution V also minimizes ‘reconstruction error’ ..

Solution: The main property that is exploited in the proof is the fact that $\text{Tr}(ABC) = \text{Tr}(BCA)$ (when dimensions are compatible). First we note that $\sum_i \|\bar{x}_i - VV^T \bar{x}_i\|^2 = \|(I - VV^T)X\|_F^2$. We will call P the projector $P = VV^T$. Then:

$$\begin{aligned}
\|(I - VV^T)X\|_F^2 &= \text{Tr}(I - P)XX^T(I - P) \\
&= \text{Tr}(XX^T - PXX^T)(I - P) \\
&= \text{Tr}(XX^T) - \text{Tr}(PXX^T) - \text{Tr}(XX^TP) + \text{Tr}(I) \\
&= \text{Tr}(XX^T) - \text{Tr}(PXX^T) - \text{Tr}(XX^TP) + \text{Tr}(I) \\
&= \text{Tr}(XX^T) - \text{Tr}(PXX^T) - \text{Tr}(XX^TP) + \text{Tr}(I) \\
&= \text{Tr}(XX^T) - \text{Tr}(PXX^T) \\
&= \text{Tr}(XX^T) - \text{Tr}(VV^TXX^T) \\
&= \text{Tr}(XX^T) - \text{Tr}(V^TXX^TV)
\end{aligned}$$

The first term is a constant, therefore the minimum is reached when the maximum of the second term is reached. \square

4 ... and that it also maximizes $\sum_{i,j} \|y_i - y_j\|_2^2$

Solution: Let us denote by \bar{y} the sample mean of the y_j s, i.e.,

$$\bar{y} = \frac{1}{n} \sum_{j=1}^n y_j.$$

We proceed backward examine the sum $\sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|_2^2$

$$\begin{aligned}\sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|_2^2 &= \sum_{i,j} \|(\mathbf{y}_i - \bar{\mathbf{y}}) - (\mathbf{y}_j - \bar{\mathbf{y}})\|_2^2 \\ &= \sum_{i,j} ((\mathbf{y}_i - \bar{\mathbf{y}}) - (\mathbf{y}_j - \bar{\mathbf{y}}), (\mathbf{y}_i - \bar{\mathbf{y}}) - (\mathbf{y}_j - \bar{\mathbf{y}})) \\ &= \sum_i \sum_j [\|(\mathbf{y}_i - \bar{\mathbf{y}})\|_2^2 + \|(\mathbf{y}_j - \bar{\mathbf{y}})\|_2^2] \dots \\ &\quad - 2 \sum_i \sum_j ((\mathbf{y}_i - \bar{\mathbf{y}}), (\mathbf{y}_j - \bar{\mathbf{y}})) \\ &= 2n \sum_i \|\mathbf{y}_i - \bar{\mathbf{y}}\|_2^2 - 2 \sum_i (\mathbf{y}_i - \bar{\mathbf{y}}, \sum_j (\bar{\mathbf{y}} - \mathbf{y}_j)) \\ &= 2n \sum_i \|\mathbf{y}_i - \bar{\mathbf{y}}\|_2^2\end{aligned}$$

The last equality comes from the fact that $\sum_j (\bar{\mathbf{y}} - \mathbf{y}_j) = \mathbf{0}$ \square

Further reading: Some references on applications of the SVD

For image processing:

<https://arxiv.org/pdf/1211.7102.pdf>

The extraordinary SVD

<https://arxiv.org/pdf/1103.2338.pdf%22%20rel=%22nofollow>

An outstanding paper for understanding the SVD and a few of its applications:

https://sites.math.washington.edu/~morrow/464_16/svd.pdf

An old paper of ours that discusses a form of truncated SVD for face recognition:

<https://www-users.cs.umn.edu/~saad/PDF/umsi-2006-16.pdf>