✍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 $\mu = \frac{1}{n}\sum x_i = \frac{1}{n}Xe$ and so,

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

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 $\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(BCA)$ (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 pojector $P = VV^T$. Then:

$$
\begin{aligned}
\|(\boldsymbol{I} - \boldsymbol{V}\boldsymbol{V}^{\boldsymbol{T}})\boldsymbol{X}\_\boldsymbol{F}^{\boldsymbol{2}}* &= \operatorname{Tr}(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{P}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}} - \boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}})(\boldsymbol{I} - \boldsymbol{P}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}) + \operatorname{Tr}(\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}) + \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}^{\boldsymbol{2}}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}) + \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{P}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{V}\boldsymbol{V}^{\boldsymbol{T}}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) \\
&= \operatorname{Tr}(\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}) - \operatorname{Tr}(\boldsymbol{V}^{\boldsymbol{T}}\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{V})
\end{aligned}
$$

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