1 Problem 1

a)
we formulate the Lagrangian function:

$$\begin{aligned}
\mathcal{L}(\theta,\lambda) &= \sum_{k=1}^{n} ||\theta - x_k||^2 + \lambda(\theta^\intercal b - 0) = \sum_{k=1}^{n}(\theta - x_k)^\intercal(\theta - x_k) + \lambda\theta^\intercal b \\
&= \sum_{k=1}^{n}(\theta^\intercal\theta + x_k^\intercal x_k - 2\theta^\intercal x_k) + \lambda\theta^\intercal b \\
&= n\theta^\intercal\theta + \sum_{k=1}^{n} x_k^\intercal x_k - 2\sum_{k=1}^{n}\theta^\intercal x_k + \lambda\theta^\intercal b \\
\frac{\partial\mathcal{L}(\theta,\lambda)}{\partial\theta} &= 2n\theta - 2\sum_{k=1}^{n} x_k + \lambda b \overset{!}{=} 0 \\
\theta^* &= \frac{2\sum\limits_{k=1}^{n} x_k + \lambda b}{2n} = \bar{x} + \frac{\lambda b}{2n}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\mathcal{L}(\theta,\lambda)}{\partial\lambda} &= \theta^\intercal b \overset{!}{=} 0 \\
\left(\bar{x} + \frac{\lambda b}{2n}\right)^\intercal b &= 0 \\
\lambda &= -2n\bar{x}^\intercal b(b^\intercal b)^{-1}
\end{aligned}$$

plug λ into the formular of θ^*:

$$\theta^* = \bar{x} - \bar{x}^\intercal b(b^\intercal b)^{-1} b$$

Geometrical interpretation:

b)

we formulate the Lagrangian function:

$$\begin{aligned}
\mathcal{L}(\theta, \lambda) &= \sum_{k=1}^{n} ||\theta - x_k||^2 + \lambda(||\theta - c||^2 - 1) = \sum_{k=1}^{n} (\theta - x_k)^\intercal(\theta - x_k) + \lambda[(\theta - c)^\intercal(\theta - c) - 1] \\
&= \sum_{k=1}^{n} (\theta^\intercal\theta + x_k^\intercal x_k - 2\theta^\intercal x_k) + \lambda(\theta^\intercal\theta + c^\intercal c - 2\theta^\intercal c - 1) \\
&= n\theta^\intercal\theta + \sum_{k=1}^{n} x_k^\intercal x_k - 2\sum_{k=1}^{n} \theta^\intercal x_k + \lambda(\theta^\intercal\theta + c^\intercal c - 2\theta^\intercal c - 1) \\
\frac{\partial \mathcal{L}(\theta, \lambda)}{\partial \theta} &= 2n\theta - 2\sum_{k=1}^{n} x_k + \lambda(2\theta - 2c) \stackrel{!}{=} 0 \\
\theta^* &= \frac{\sum_{k=1}^{n} x_k + \lambda c}{n + \lambda} \\
\frac{\partial \mathcal{L}(\theta, \lambda)}{\partial \lambda} &= \theta^\intercal\theta + c^\intercal c - 2\theta^\intercal c - 1 \stackrel{!}{=} 0
\end{aligned}$$

Geometrical interpretation:

2 Problem 2

a)

We know that if A is an $n \times n$ matrix and let λ_1,..,λ_n be its eigenvalues. Here $det(A)$ is the determinant of the matrix A and $tr(A)$ is the trace of the matrix A, and the determinant of A is the product of its eigenvalues, and the trace of A is the sum of the eigenvalues. Hence, the trace of S equals to the sum of eigenvalues, and $S_i i$ is the diagonal elements of the scatter matrix, therefore:

$$tr(S) = \sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} S_{ii}$$

It holds that $\lambda_i \geq 0, \forall i = 1, ..., d$, since S is positive semi-definite,

$$\lambda_1 \leq \sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} S_{ii}$$

b)

It holds that $\lambda_1 = \sum_{i=1}^{d} S_{ii}$ if and only if $\lambda_i = 0, \forall i = 2, ..., d$. In this case the matrix S is of rank 1, which means that all features are linearly dependent.

3 Problem 3

a)
we formulate the formulas from question:

$$
\begin{aligned}
J(\mathbf{w}) &= ||\mathbf{S}\mathbf{w}|| - \frac{1}{2}\mathbf{w}^\intercal \mathbf{S}\mathbf{w} \\
\mathbf{v} &= \mathbf{S}^{\frac{1}{2}}\mathbf{w} \qquad \mathbf{w} = \mathbf{S}^{-\frac{1}{2}}\mathbf{v} \\
J(\mathbf{w}) &= ||\mathbf{S}\mathbf{S}^{-\frac{1}{2}}\mathbf{v}|| - \frac{1}{2}(\mathbf{S}^{-\frac{1}{2}}\mathbf{v})^\intercal \mathbf{S}\mathbf{S}^{-\frac{1}{2}}\mathbf{v} \\
&= ||\mathbf{S}^{\frac{1}{2}}\mathbf{v}|| - \frac{1}{2}\mathbf{v}^\intercal\mathbf{v} \\
&= ||\mathbf{S}^{\frac{1}{2}}\mathbf{v}|| - \frac{1}{2}||\mathbf{v}||^2 \\
\frac{\partial J(\mathbf{v})}{\partial \mathbf{v}} &= \frac{\mathbf{S}^{\frac{1}{2}}\mathbf{v}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{v}||}\mathbf{S}^{\frac{1}{2}} - \mathbf{v} \\
\mathbf{v} &\leftarrow \mathbf{v} + \gamma\frac{\partial J}{\partial \mathbf{v}} = \mathbf{v} + \gamma\frac{\mathbf{S}^{\frac{1}{2}}\mathbf{v}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{v}||}\mathbf{S}^{\frac{1}{2}} - \mathbf{v} \\
\mathbf{v} &\leftarrow \gamma\frac{\mathbf{S}\mathbf{v}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{v}||} \\
&\text{If } \gamma \text{ is identity matrix} \\
\mathbf{S}^{-\frac{1}{2}}\mathbf{v} &\leftarrow \mathbf{S}^{-\frac{1}{2}}\frac{\mathbf{S}\mathbf{v}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{v}||} = \frac{\mathbf{S}^{\frac{1}{2}}\mathbf{v}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{v}||} \\
\mathbf{w} &\leftarrow \frac{\mathbf{S}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{w}}{||\mathbf{S}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{w}||} \\
\mathbf{w} &\leftarrow \frac{\mathbf{S}\mathbf{w}}{||\mathbf{S}\mathbf{w}||}
\end{aligned}
$$

b)

$$||\mathbf{w}|| = \left|\left|\frac{\mathbf{S}\mathbf{w}}{||\mathbf{S}\mathbf{w}||}\right|\right| = \frac{||\mathbf{S}\mathbf{w}||}{||\mathbf{S}\mathbf{w}||} = 1$$