

Machine Learning 2

Exercise Sheet 1

Exercise 1

(i)

Given the following problem:

$$\begin{aligned} \min_w E(w) &= \sum_i \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } &\sum_j w_{ij} = 1 \end{aligned}$$

we are trying to prove that the multiplication of each vector \vec{x}_i by a constant scalar $\alpha \in \mathbb{R}^+ \setminus \{0\}$ does not alter the problem's solution.

$$\begin{aligned} &\min_w \sum_i \left| \alpha \vec{x}_i - \sum_j w_{ij} \alpha \vec{x}_j \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \sum_i \left| \alpha \vec{x}_i - \alpha \sum_j w_{ij} \vec{x}_j \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \sum_i \alpha^2 \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \alpha^2 \sum_i \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \end{aligned}$$

Since the multiplication by $\alpha^2 \in \mathbb{R}^+ \setminus \{0\}$ doesn't change the minima with respect to w , the minimization problem remains the same

(ii)

Given the following problem:

$$\begin{aligned} \min_w E(w) &= \sum_i \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } &\sum_j w_{ij} = 1 \end{aligned}$$

we are trying to prove that the addition of constant vector $\vec{v} \in \mathbb{R}^D$ to each vector \vec{x}_i does not alter the problem's solution.

$$\begin{aligned} &\min_w \sum_i \left| (x_i + \vec{v}) - \sum_j w_{ij} (x_j + \vec{v}) \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \sum_i \left| x_i + \vec{v} - \left(\sum_j w_{ij} x_j + \sum_j w_{ij} \vec{v} \right) \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \sum_i \left| x_i + \vec{v} - \left(\sum_j w_{ij} x_j \right) - \left(\sum_j w_{ij} \vec{v} \right) \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \text{minimization constraint} \\ &\quad \equiv \\ &\min_w \sum_i \left| x_i + \vec{v} - \left(\sum_j w_{ij} x_j \right) - \vec{v} \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \\ &\quad \equiv \\ &\min_w \sum_i \left| x_i - \sum_j w_{ij} x_j \right|^2 \\ &\quad \text{s. t. } \sum_j w_{ij} = 1 \end{aligned}$$

(iii)

Given the following problem:

$$\begin{aligned} \min_w E(w) &= \sum_i \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } &\sum_j w_{ij} = 1 \end{aligned}$$

we are trying to prove that the multiplication of each vector \vec{x}_i by a constant, orthogonal $D \times D$ matrix U does not alter the problem's solution.

$$\begin{aligned} \min_w \sum_i \left| U\vec{x}_i - \sum_j w_{ij} U\vec{x}_j \right|^2 \\ \text{s. t. } \sum_j w_{ij} &= 1 \\ &\equiv \\ \min_w \sum_i \left| U\vec{x}_i - \sum_j U w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } \sum_j w_{ij} &= 1 \\ &\equiv \\ \min_w \sum_i \left| U\vec{x}_i - U \sum_j w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } \sum_j w_{ij} &= 1 \\ &\equiv \\ \min_w \sum_i \left| U \left(\vec{x}_i - \sum_j w_{ij} \vec{x}_j \right) \right|^2 \\ \text{s. t. } \sum_j w_{ij} &= 1 \end{aligned}$$

Since for all orthogonal matrices $U \in \mathbb{R}^{D \times D}$ and vectors $\vec{x} \in \mathbb{R}^D$ we have that $|U\vec{x}| = |\vec{x}|$

$$\begin{aligned} &\equiv \\ \min_w \sum_i \left| \vec{x}_i - \sum_j w_{ij} \vec{x}_j \right|^2 \\ \text{s. t. } \sum_j w_{ij} &= 1 \end{aligned}$$

Exercise 2

(i)

We have to prove $w^\top C w \stackrel{?}{=} \epsilon \stackrel{(3)}{=} \|x - \sum_j w_j \eta_j\|^2$.

$$\begin{aligned}
& w^\top C w \\
&= w^\top (\mathbb{1}_k x^\top - \eta)(\mathbb{1}_k x^\top - \eta)^\top w \\
&= w^\top (\mathbb{1}_k x^\top - \eta)(x \mathbb{1}_k^\top - \eta^\top) w \\
&= w^\top \mathbb{1}_k x^\top x \mathbb{1}_k^\top w - w^\top \mathbb{1}_k x^\top \eta^\top w - w^\top \eta x \mathbb{1}_k^\top w + w^\top \eta \eta^\top w \\
&= (w^\top \mathbb{1}_k x^\top)(x \mathbb{1}_k^\top w) - 2(w^\top \mathbb{1}_k x^\top)(\eta^\top w) + (w^\top \eta)(\eta^\top w) \\
&= \|(w^\top \mathbb{1}_k x^\top) - (w^\top \eta)\|^2 \\
&= \|w^\top (\mathbb{1}_k x^\top - \eta)\|^2 \\
&= \left\| \sum_j w_j (x - \eta_j) \right\|^2
\end{aligned}$$

Since $\sum_i w_i = 1$ we find $\sum_i w_i x = x$, which leads to the desired result. \square

We now perform Lagrange optimization:

$$\begin{aligned}
\Lambda(w, \lambda) &= w^\top C w - \lambda(w^\top \mathbb{1}_k - 1) \\
\frac{\partial \Lambda}{\partial w} &= 2Cw - \lambda \mathbb{1}_k \stackrel{!}{=} 0 \quad \implies \quad 2Cw = \lambda \mathbb{1}_k \quad \implies \quad w = \frac{\lambda}{2} C^{-1} \mathbb{1}_k \\
\frac{\partial \Lambda}{\partial \lambda} &= w^\top \mathbb{1}_k - 1 \stackrel{!}{=} 0 \quad \implies \quad w^\top \mathbb{1}_k = 1
\end{aligned}$$

We now replace w in the constraint $w^\top \mathbb{1}_k = 1$. Since $C = C^\top$ we find

$$\frac{\lambda}{2} = \frac{1}{\mathbb{1}_k^\top C^{-1} \mathbb{1}_k}$$

(ii)

Replacing $\lambda/2$ in the deduced definition of w leads to the desired result. The candidate w is indeed a minimum since $\frac{\partial^2 \Lambda}{\partial w^2} = 2C$ (invertible covariance matrices have positive definite quadratic forms). \square