Optimization Models in Engineering—Homework 2

1. Consider the subspace $\mathcal{S} = \text{span}(x^{(1)}, x^{(2)}, x^{(3)})$, where

$$x^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad x^{(2)} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad x^{(3)} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \tag{1}$$

   i) Find the dimension of $\mathcal{S}$.

   ii) Calculate the projection of the point $y = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ on $\mathcal{S}$.

Solution.

i) Note that $x^{(1)}$ and $x^{(2)}$ are linearly independent by inspection. Furthermore, $x^{(3)} = x^{(1)} + x^{(2)}$. Therefore $\mathcal{S}$ is spanned by $x^{(1)}$ and $x^{(2)}$ and is of dimension 2.

ii) We want to find a point $y^* \in \text{span}(x^{(1)}, x^{(2)})$ such that $y - y^*$ is orthogonal to both $x^{(1)}$ and $x^{(2)}$. We can write the orthogonality condition out as:

$$\begin{aligned} 1(1 - y_1^*) + 1(2 - y_2^*) + 1(4 - y_3^*) &= 0 \\ -1(1 - y_1^*) + 0(2 - y_2^*) + 1(4 - y_3^*) &= 0 \end{aligned}$$

Simplifying, we get the conditions:

$$\begin{aligned} y_1^* + y_2^* + y_3^* &= 7 \\ y_1^* - y_3^* &= -3 \end{aligned}$$

Now since $y^* \in \text{span}(x^{(1)}, x^{(2)})$, we can write

$$y^* = \alpha x^{(1)} + \beta x^{(2)} = (\alpha - \beta, \alpha, \alpha + \beta),$$

with $\alpha, \beta \in \mathbb{R}$. Substituting into our orthogonality conditions yields:

$$\begin{aligned} (\alpha - \beta) + \alpha + \alpha + \beta &= 7, \\ (\alpha - \beta) - (\alpha + \beta) &= -3. \end{aligned}$$

The second equation gives us that $\beta = \frac{3}{2}$, and the first equation yields $\alpha = \frac{7}{3}$. Finally, we get:

$$y^* = \frac{7}{3} x^{(1)} + \frac{3}{2} x^{(2)} = \left(\frac{5}{6}, \frac{7}{3}, \frac{23}{6} \right) \approx (0.833, 2.33, 3.83).$$

2. Consider the box $\mathcal{S}_1$ and ball $\mathcal{S}_2$ defined as

$$\mathcal{S}_1 = \{x \in \mathbb{R}^2 \mid -2 \leq x_1 \leq 2,\ -0.5 \leq x_2 \leq 0.5\}, \quad \mathcal{S}_2 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\} \tag{2}$$

Given a point $z \in \mathbb{R}^2$, write an optimization problem in a standard form that finds the projection of z onto the set $\mathcal{S}_1 \cap \mathcal{S}_2$ (i.e., the solution of the optimization problem should correspond to the closest point in $\mathcal{S}_1 \cap \mathcal{S}_2$ to z; note that you do not need to solve the optimization problem).

Solution.

We can write our optimization problem as follows:

$$\begin{aligned} \operatorname*{arg\,min}_{x \in \mathbb{R}^2} \quad & \|z - x\|_2^2 \\ x_1 - 2 &\leq 0 \\ -x_1 - 2 &\leq 0 \\ x_2 - 0.5 &\leq 0 \\ -x_2 - 0.5 &\leq 0 \\ x_1^2 + x_2^2 - 1 &\leq 0 \end{aligned}$$

3. A company has n factories. Factory i (for $i = 1, 2, ..., n$) is located at point (a_i, b_i) in the two-dimensional plane $\mathbb{R}^2$. The company wants to locate a warehouse at a point (x_1, x_2) that minimizes

$$\sum_{i=1}^{n} (\text{distance from factory } i \text{ to the warehouse})^2$$

Find all possible values of the point (x_1^*, x_2^*) that satisfy the necessary condition for local optimality.

Solution.

Formally, we are attempting to minimize the objective:

$$\min_{x_1, x_2} \quad \sum_{i=1}^{n} (a_i - x_1)^2 + (b_i - x_2)^2.$$

Taking the gradient of the objective with respect to each variable gives us

$$\frac{d}{dx_1} \sum_{i=1}^{n} (a_i - x_1)^2 + (b_i - x_2)^2 = -\sum_{i=1}^{n} 2(a_i - x_1),$$
$$\frac{d}{dx_2} \sum_{i=1}^{n} (a_i - x_1)^2 + (b_i - x_2)^2 = -\sum_{i=1}^{n} 2(b_i - x_2).$$

According to the necessary optimality condition, the optimal x_1^* and x_2^* must make the above derivatives equal to zero. Rearranging, we get the conditions

$$\sum_{i=1}^{n} x_1^* = n x_1^* = \sum_{i=1}^{n} a_i, \qquad \sum_{i=1}^{n} x_2^* = n x_2^* = \sum_{i=1}^{n} b_i,$$

and hence,

$$x_1^* = \frac{1}{n} \sum_{i=1}^{n} a_i, \qquad x_2^* = \frac{1}{n} \sum_{i=1}^{n} b_i.$$

So there is only one solution: set the warehouse at the average position of the factories.

4. Given a natural number $k \in \{1, 2, \dots\}$, a symmetric matrix $P \in \mathbb{R}^{n \times n}$, a vector $q \in \mathbb{R}^n$ and a scalar $r \in \mathbb{R}$, consider the optimization problem

$$\min_{x \in \mathbb{R}^n} \quad (x^\top P x)^k + q^\top x + r \tag{3}$$

Assume that q is a nonzero vector.

i) Calculate the gradient of the function $q^\top x$.

ii) Calculate the gradient of the function $x^\top P x$.

iii) Calculate the gradient of the objective function of the optimization problem (3).

iv) Given a point x^*, write the necessary optimality condition for x^* to be a local minimum of the optimization problem (3).

v) Assume that q is not in the range of P. Prove that the optimization problem (3) cannot have any local minimum (hint: show that the necessary optimality condition has no solution).

vi) Assume that P is invertible. Given a local minimum x^* of the optimization problem (3), show that there is a scalar α such that $x^* = \alpha P^{-1} q$.

vii) Again assume that P is invertible. Solve for α in Part (vi) and calculate it in terms of only the known parameters P, q, r, k (hint: Substitute the formula $x^* = \alpha P^{-1} q$ into the optimality condition and write it in terms of α).

Solution.

i) I'm going to use t to index x since we have a lot of summations going on. Note that $q^\top x = \sum_i q_i x_i$. Therefore

$$\frac{d}{dx_t} q^\top x = q_t,$$

and $\nabla_x q^\top x = q$.

ii) Now we have

$$\begin{aligned} x^\top P x &= \sum_{i=1}^{n} \sum_{j=1}^{n} P_{i,j} x_i x_j \\ &= \sum_{i=1}^{n} P_{i,i} x_i^2 + \sum_{i=1}^{n} \sum_{j \neq i} P_{i,j} x_i x_j. \end{aligned}$$

Differentiating gives us

$$\begin{aligned} \frac{d}{dx_t} x^\top P x &= 2P_{t,t} x_t + \sum_{j \neq t} P_{t,j} x_j + \sum_{i \neq t} P_{i,t} x_i \\ &= 2P_{t,t} x_t + 2 \sum_{j \neq t} P_{t,j} x_j \qquad \text{due to symmetry of } P \\ &= 2 \sum_{j=1}^{n} P_{t,j} x_j. \end{aligned}$$

Now we conclude by combining over t:

$$\nabla_x (x^\top P x) = 2Px.$$

iii) Apply the chain rule to the first term gives us:

$$\begin{aligned} \nabla_x (x^\top P x)^k &= k (x^\top P x)^{k-1} \left(\nabla_x (x^\top P x) \right) \\ &= 2k (x^\top P x)^{k-1} P x \end{aligned}$$

Given that the gradient of the scalar r with respect to x is zero, we can write out the gradient of the objective function as

$$\nabla_x (x^\top P x)^k + q^\top x + r = 2k(x^\top P x)^{k-1} P x + q$$

iv) We just want that the gradient at x^* is equal to zero for optimality:

$$2k({x^*}^\top P x^*)^{k-1} P x^* = -q$$

v) Assume for the sake of contradiction that there existed a solution x^* to the optimality condition. Note that $({x^*}^\top P x^*)^{k-1}$ is a scalar; we can thus observe that the optimality condition takes the form

$$cPx^* = -q$$

where $c = 2k({x^*}^\top P x^*)^{k-1} \in \mathbb{R}$. By the problem assumption that $q \neq 0$, we have that $c \neq 0$. However, this statement precisely implies that q is in the range of P, and we have a contradiction.

vi) Any local minimum must satisfy the optimality condition

$$cPx^* = -q,$$

where c is the same as in the previous part. If $c = 0$, then $q = 0$, which violates the problem assumption. Thus $c \neq 0$. Inverting P and rearranging, this implies that:

$$x^* = -c^{-1} P^{-1} q = \alpha P^{-1} q,$$

with $\alpha = -c^{-1}$.

vii) Recall that the optimality condition is

$$2k({x^*}^\top P x^*)^{k-1} P x^* = -q$$

Substituting into this gets us:

$$\begin{aligned}
2k((\alpha P^{-1} q)^\top P (\alpha P^{-1} q))^{k-1} P(\alpha P^{-1} q) &= -q \\
2k\alpha^{2(k-1)+1}((P^{-1} q)^\top P (P^{-1} q))^{k-1} q &= -q \\
2k\alpha^{2k-1}(q^\top P^{-1} q)^{k-1} q &= -q
\end{aligned}$$

This is satisfied if

$$\begin{aligned}
2k\alpha^{2k-1}(q^\top P^{-1} q)^{k-1} &= -1 \\
\alpha^{2k-1} &= -(2k)^{-1} \cdot (q^\top P^{-1} q)^{1-k} \\
\alpha &= -(2k)^{\frac{-1}{2k-1}} \cdot (q^\top P^{-1} q)^{\frac{1-k}{2k-1}}
\end{aligned}$$