Optimization Models: Homework #2

Due on September 23, 2025 at 11:59pm

Zachary Brandt

Problem 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 $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$ on $\mathcal{S}$.

Solution

The dimension of the subspace $\mathcal{S}$ is 2, because the span is not linearly independent. Vector $x^{(3)}$ is within the span of $x^{(1)}$ and $x^{(2)}$

$$x^{(1)} + x^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = x^{(3)}.$$

The other two vectors span a plane with dimension 2. To calculate the projection of the point $x = (2, 3, 5)$ onto $\mathcal{S}$, I solve a system of equations to find $x^* = \alpha_1 x^{(1)} + \alpha_2 x^{(2)}$. The vector pointing between x and its projection onto the span should be normal to $x^{(1)}$ and $x^{(2)}$:

$$\begin{aligned}
\langle x - x^*, x^{(1)} \rangle &= \langle x - \alpha_1 x^{(1)} - \alpha_2 x^{(2)}, x^{(1)} \rangle = 0 \\
\langle x - x^*, x^{(2)} \rangle &= \langle x - \alpha_1 x^{(1)} - \alpha_2 x^{(2)}, x^{(2)} \rangle = 0 \\
\Rightarrow \langle x, x^{(1)} \rangle &= \alpha_1 \langle x^{(1)}, x^{(1)} \rangle + \alpha_2 \langle x^{(2)}, x^{(1)} \rangle \\
\Rightarrow \langle x, x^{(2)} \rangle &= \alpha_1 \langle x^{(1)}, x^{(2)} \rangle + \alpha_2 \langle x^{(2)}, x^{(2)} \rangle
\end{aligned}$$

which then equals

$$\begin{aligned}
10 &= \alpha_1 \cdot 3 + \alpha_2 \cdot 0 \\
3 &= \alpha_1 \cdot 0 + \alpha_2 \cdot 2
\end{aligned}$$

Solving this system of equations gives $\alpha_1 = \frac{10}{3}$ and $\alpha_2 = \frac{3}{2}$. Therefore, the projection of the point x onto $\mathcal{S}$ is given by

$$x^* = \frac{10}{3} x^{(1)} + \frac{3}{2} x^{(2)} = \left(\frac{11}{6}, \frac{10}{3}, \frac{29}{6} \right).$$

Problem 2

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

$$\mathcal{S}_1 = \{x \in \mathbb{R}^2 \mid -3 \le x_1 \le 3,\ -1 \le x_2 \le 1\}, \quad \mathcal{S}_2 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \le 2\} \tag{2}$$

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

$$\begin{aligned}
\arg\min_{\boldsymbol{x}} \quad & \underbrace{\boxed{\|\boldsymbol{x} - \boldsymbol{z}\|_2}}_{f_0(\boldsymbol{x})}, \\
\text{s.t.} \quad & \underbrace{\boxed{x_1 - 3}}_{f_1(\boldsymbol{x})} \le 0, \\
& \underbrace{\boxed{-x_1 - 3}}_{f_2(\boldsymbol{x})} \le 0, \\
& \underbrace{\boxed{x_2 - 1}}_{f_3(\boldsymbol{x})} \le 0, \\
& \underbrace{\boxed{-x_2 - 1}}_{f_4(\boldsymbol{x})} \le 0, \\
& \underbrace{\boxed{x_1^2 + x_2^2 - 2}}_{f_5(\boldsymbol{x})} \le 0.
\end{aligned}$$

where

- $(x_1, x_2) \in \mathbb{R}^2$ is the *decision variable*;
- $f_0 : \mathbb{R}^2 \to \mathbb{R}$ is the *objective function*, or the *Euclidean distance* between x and z;
- f_1 & $f_2 : \mathbb{R}^2 \to \mathbb{R}$ represent the left and right bounds on the box $\mathcal{S}_1$;
- f_3 & $f_4 : \mathbb{R}^2 \to \mathbb{R}$ represent the upper and lower bounds on the box $\mathcal{S}_1$;
- $f_5 : \mathbb{R}^2 \to \mathbb{R}$ represents the constraint on the ball $\mathcal{S}_2$.

Problem 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

A function for the sum of the distances from each factory to the warehouse is:

$$f(x_1, x_2) = \sum_{i=1}^{n} \left(\sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2} \right).$$

Since the square root of the function is only increasing, the arguments of the minimum of the function without the square root will be the same as the arguments of the minimum of the function with the square root. Therefore, I will minimize the following function:

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

The gradient of $g(x_1, x_2)$ is a vector of the two partial derivatives:

$$\nabla g(x_1, x_2) = \begin{bmatrix} \frac{\partial g(x_1, x_2)}{\partial x_1} \\ \frac{\partial g(x_1, x_2)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2\sum_{i=1}^{n}(x_1 - a_i) \\ 2\sum_{i=1}^{n}(x_2 - b_i) \end{bmatrix} = \begin{bmatrix} 2nx_1 - 2\sum_{i=1}^{n} a_i \\ 2nx_2 - 2\sum_{i=1}^{n} b_i \end{bmatrix}.$$

Setting these equal to zero gives the necessary conditions for local optimality:

$$\begin{aligned} 2nx_1 - 2\sum_{i=1}^{n} a_i = 0 \Rightarrow x_1 = \frac{1}{n}\sum_{i=1}^{n} a_i \\ 2nx_2 - 2\sum_{i=1}^{n} b_i = 0 \Rightarrow x_2 = \frac{1}{n}\sum_{i=1}^{n} b_i \end{aligned}.$$

So, x_1 and x_2 are the averages of the a_i's and b_i's.

Problem 4

Given a natural number $k \in \{1, 2, \ldots\}$, 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

The gradient of the function $q^\top x$ is the vector q itself:

$$\nabla(q^\top x) = q.$$

I found the gradient of the function $x^\top P x$ by expressing the function in sum notation and then considering the partial derivative with respect to any component x_i of the vector x:

$$x^\top P x = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i \cdot P_{ij} x_j) .$$

Taking the partial derivative with respect to x_1, for example, gives

$$\begin{aligned}
\frac{\partial}{\partial x_1}(x^\top P x) &= \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial}{\partial x_1}(x_j \cdot P_{jk} x_k) \\
&= \sum_{j=1}^{n} P_{1j} x_j + \sum_{i=1}^{n} P_{i1} x_i \\
&= (Px)_1 + (P^\top x)_1 \\
&= 2(Px)_1,
\end{aligned}$$

since P is symmetric. Repeating this for each component of x gives the gradient

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

Problem 4 continued on next page. . .

The gradient of the objective function is then the following:

$$\begin{aligned}\nabla\left((x^\top Px)^k + q^\top x + r\right) &= \nabla\left((x^\top Px)^k\right) + \nabla(q^\top x) + \nabla r \\ &= k(x^\top Px)^{k-1}\nabla(x^\top Px) + q + 0 \\ &= k(x^\top Px)^{k-1}(2Px) + q \\ &= 2k(x^\top Px)^{k-1}Px + q.\end{aligned}$$

The necessary condition for x^* to be a local minimum is that the gradient of the objective function is zero at that point, when $x = x^*$,

$$2k(x^\top Px)^{k-1}Px + q = 0.$$

Now if q is not in the range of P, then there is no solution. I rearrange the equation to isolate q on one side to make this clear:

$$q = -2k(x^\top Px)^{k-1}Px.$$

The right-hand side is in the range of P for any x, since everything before the P is a scalar multiple of x. That's why if q is not in the range of P, there is no solution to the equation.
Now assume that P is invertible. Then, given a local minimum x^*, I can rearrange the necessary condition to isolate x^*:

$$\begin{aligned}2k(x^{*\top}Px^*)^{k-1}Px^* + q &= 0 \\ 2k(x^{*\top}Px^*)^{k-1}Px^* &= -q \\ Px^* &= -\frac{1}{2k(x^{*\top}Px^*)^{k-1}}q \\ x^* &= -\frac{1}{2k(x^{*\top}Px^*)^{k-1}}P^{-1}q.\end{aligned}$$

So letting $\alpha = -\frac{1}{2k(x^{*\top}Px^*)^{k-1}}$ gives the desired result.
Finally, I substitute $x^* = \alpha P^{-1}q$ into the necessary condition and solve for α:

$$\begin{aligned}2k(x^{*\top}Px^*)^{k-1}Px^* + q &= 0 \\ 2k((\alpha P^{-1}q)^\top P(\alpha P^{-1}q))^{k-1}P(\alpha P^{-1}q) + q &= 0 \\ 2k(\alpha^2 q^\top P^{-1}q)^{k-1}\alpha q + q &= 0 \\ 2k\alpha^{2k-1}(q^\top P^{-1}q)^{k-1}q + q &= 0 \\ (2k\alpha^{2k-1}(q^\top P^{-1}q)^{k-1} + 1)q &= 0.\end{aligned}$$

Since q is nonzero, the term in parentheses must be zero:

$$\begin{aligned}2k\alpha^{2k-1}(q^\top P^{-1}q)^{k-1} + 1 &= 0 \\ 2k\alpha^{2k-1}(q^\top P^{-1}q)^{k-1} &= -1 \\ \alpha^{2k-1} &= -\frac{1}{2k(q^\top P^{-1}q)^{k-1}} \\ \alpha &= \left(-\frac{1}{2k(q^\top P^{-1}q)^{k-1}}\right)^{\frac{1}{2k-1}}.\end{aligned}$$

This gives α in terms of only the known parameters P, q, k. Note that r does not appear in the expression for α, so it is not needed.