EECS 127: Optimization Models in Engineering
Homework 2

Your Name

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

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

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

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

Solution 1. *(i) Dimension of $\mathcal{S}$:*

To find the dimension of $\mathcal{S}$, we'll perform Gaussian elimination on the matrix $\mathbf{A}$ formed by the given vectors:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix} \quad \text{(R2} \to \text{R2 - R1, R3} \to \text{R3 - R1)}$$

$$\sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(R3} \to \text{R3 - 2R2)}$$

The rank of the matrix is equal to the number of non-zero rows in this row echelon form, which is 2. Therefore, $\dim(\mathcal{S}) = 2$.

(ii) Projection onto $\mathcal{S}$:

Let $\mathbf{b} := \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$. We'll use the Gram-Schmidt process to find an orthonormal basis for $\mathcal{S}$, then project $\mathbf{b}$ onto this basis.

1. Apply Gram-Schmidt process:

$$\mathbf{u}_1 = \mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{u}_2 = \mathbf{x}^{(2)} - \text{proj}_{\mathbf{u}_1}(\mathbf{x}^{(2)}) = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - 0\mathbf{u}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\mathbf{u}_3 = \mathbf{x}^{(3)} - \text{proj}_{\mathbf{u}_1}(\mathbf{x}^{(3)}) - \text{proj}_{\mathbf{u}_2}(\mathbf{x}^{(3)}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

2. Normalize the basis vectors:

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

3. Calculate the projection:

$$\begin{aligned} \text{proj}_{\mathcal{S}}(\mathbf{b}) &= \langle \mathbf{b}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{b}, \mathbf{v}_2 \rangle \mathbf{v}_2 \\ &= \frac{7}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \\ &= \frac{7}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{7}{3} - \frac{3}{2} \\ \frac{7}{3} \\ \frac{7}{3} + \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ \frac{7}{3} \\ \frac{17}{6} \end{pmatrix} \end{aligned}$$

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

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

Given a point $\mathbf{z} \in \mathbb{R}^2$, write an optimization problem in standard form that finds the projection of $\mathbf{z}$ onto the set $\mathcal{S}_1 \cap \mathcal{S}_2$.

Solution 2. The optimization problem can be formulated as:

$$\begin{aligned} \min_{\mathbf{z}^*} \quad & \|\mathbf{z} - \mathbf{z}^*\|_2^2 \\ \text{s.t.} \quad & (z_1^*)^2 + (z_2^*)^2 \le 1 \\ & -2 \le z_1^* \le 2 \\ & -0.5 \le z_2^* \le 0.5 \end{aligned}$$

Note: We use the squared norm to simplify calculations, as it doesn't change the optimal solution.

Problem 3. A company has n factories. Factory i (for $i = 1, 2, \ldots, n$) 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 3. Let's formulate the minimization above as follows:

$$\sum_{i=1}^{n} (\text{distance from factory } i \text{ to the warehouse})^2 = \sum_{i=1}^{n} \left[(a_i - x_1)^2 + (b_i - x_2)^2\right]$$

We need to solve for the necessary condition of local optimality where

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2nx_1 - 2\sum_{i=1}^{n} a_i \\ 2nx_2 - 2\sum_{i=1}^{n} b_i \end{pmatrix} = \mathbf{0}$$

This means

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} a_i \\ \frac{1}{n}\sum_{i=1}^{n} b_i \end{pmatrix}$$

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

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

Assume that $\mathbf{q}$ is a nonzero vector.

(i) Calculate the gradient of the function $\mathbf{q}^\top \mathbf{x}$.

(ii) Calculate the gradient of the function $\mathbf{x}^\top \mathbf{P} \mathbf{x}$.

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

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

(v) Assume that $\mathbf{q}$ is not in the range of $\mathbf{P}$. Prove that the optimization problem cannot have any local minimum.

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

(vii) Again assume that $\mathbf{P}$ is invertible. Solve for α in Part (vi) and calculate it in terms of only the known parameters $\mathbf{P}, \mathbf{q}, r, k$.

Solution 4. (i) $\nabla \mathbf{q}^\top \mathbf{x} = \mathbf{q}$

(ii) $\nabla \mathbf{x}^\top \mathbf{P} \mathbf{x}$

Step 1: Expand the quadratic form

First, let's write out what $\mathbf{x}^\top \mathbf{P} \mathbf{x}$ means: $\mathbf{x}^\top \mathbf{P} \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i p_{ij} x_j$ where p_{ij} is the element in the i-th row and j-th column of P.

Step 2: Calculate the partial derivatives
To find the gradient, we need to calculate the partial derivative with respect to each x_k: $\frac{\partial}{\partial x_k}(\mathbf{x}^\top \mathbf{P}\mathbf{x}) = \frac{\partial}{\partial x_k} \sum_{i=1}^n \sum_{j=1}^n x_i p_{ij} x_j$
Using the product rule and noting that p_{ij} is constant: $\frac{\partial}{\partial x_k}(\mathbf{x}^\top \mathbf{P}\mathbf{x}) = \sum_{i=1}^n p_{ik} x_i + \sum_{j=1}^n p_{kj} x_j$

Step 3: Simplify using matrix notation
We can write this more compactly as: $\frac{\partial}{\partial x_k}(\mathbf{x}^\top \mathbf{P}\mathbf{x}) = (\mathbf{P}\mathbf{x})_k + (\mathbf{P}^\top \mathbf{x})_k$ where $(\mathbf{P}\mathbf{x})_k$ denotes the k-th element of the vector $\mathbf{P}\mathbf{x}$.

Step 4: Combine results for all k
Combining the results for all k, we get the gradient vector: $\nabla \mathbf{x}^\top \mathbf{P}\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{P}^\top \mathbf{x} = 2\mathbf{P}\mathbf{x}$ as P is symmetric.

(iii) $\nabla(\mathbf{x}^\top \mathbf{P}\mathbf{x})^k + \mathbf{q}^\top \mathbf{x} + r = 2k(\mathbf{x}^\top \mathbf{P}\mathbf{x})^{k-1}\mathbf{P}\mathbf{x} + \mathbf{q}$
(iv) $2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}\mathbf{P}\mathbf{x}^* + \mathbf{q} = \mathbf{0}$
(v)
Proof that no local minimum exists when $\mathbf{q}$ *is not in the range of* $\mathbf{P}$*:*
Assume, for contradiction, that the problem has a local minimum. This means the necessary optimality condition has a solution $\mathbf{x}^*$ such that:

$$2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}\mathbf{P}\mathbf{x}^* + \mathbf{q} = \mathbf{0}$$
$$-2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}\mathbf{P}\mathbf{x}^* = \mathbf{q}$$

This implies that $\mathbf{q}$ is in the range of $\mathbf{P}$, contradicting our assumption. Therefore, the optimization problem cannot have any local minimum. □
(vi)
Showing $\mathbf{x}^* = \alpha\mathbf{P}^{-1}\mathbf{q}$ *for some scalar* α*:*
From the necessary optimality condition:

$$2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}\mathbf{P}\mathbf{x}^* + \mathbf{q} = \mathbf{0}$$
$$\mathbf{P}\mathbf{x}^* = -\frac{1}{2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}}\mathbf{q}$$

Since $\mathbf{P}$ is invertible, we can multiply both sides by $\mathbf{P}^{-1}$:

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

Let $\alpha = -\frac{1}{2k(\mathbf{x}^{*\top}\mathbf{P}\mathbf{x}^*)^{k-1}}$. Then $\mathbf{x}^* = \alpha\mathbf{P}^{-1}\mathbf{q}$.
(vii)
Solving for α*:*
Substitute $\mathbf{x}^* = \alpha\mathbf{P}^{-1}\mathbf{q}$ into the optimality condition:

$$2k(\alpha^2\mathbf{q}^\top\mathbf{P}\mathbf{P}^{-1}\mathbf{q})^{k-1}\mathbf{P}\alpha\mathbf{P}^{-1}\mathbf{q} + \mathbf{q} = \mathbf{0}$$
$$2k\alpha^{2k-1}\mathbf{P}\mathbf{q} + \mathbf{q} = \mathbf{0}$$
$$2k\alpha^{2k-1}\mathbf{q} + \mathbf{q} = \mathbf{0}$$
$$(2k\alpha^{2k-1} + 1)\mathbf{q} = \mathbf{0}$$

Since $\mathbf{q}$ is nonzero, we have $2k\alpha^{2k-1} + 1 = 0$. Solving for α:

$$\alpha^{2k-1} = -\frac{1}{2k}$$
$$\alpha = \left(-\frac{1}{2k}\right)^{\frac{1}{2k-1}}$$