Discussion 12

1. Median versus average

This question illustrates the connection between the median and the mean of a finite set of real numbers.

For a given vector $\vec{v} \in \mathbb{R}^n$, the average can be found as the solution to the optimization problem

$$\min_{x \in \mathbb{R}} \|\vec{v} - x\vec{1}\|_2^2, \tag{1}$$

where $\vec{1}$ denotes the vector of ones in $\mathbb{R}^n$. Similarly, it turns out that the median (a median is any value x such that it is possible to partition the values at x (if any) as being either just below x, at x, or just above x, in such a way that there is an equal number of values in $\vec{v}$ above and below x) can be found via the optimization problem

$$\min_{x \in \mathbb{R}} \|\vec{v} - x\vec{1}\|_1. \tag{2}$$

We consider a robust version of problem (1) of finding the average, i.e.

$$\min_{x} \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq \lambda} \|\vec{v} + \vec{u} - x\vec{1}\|_2^2, \tag{3}$$

in which we assume that the components of $\vec{v}$ can be independently perturbed by a vector $\vec{u}$ each of whose components has magnitude bounded by a given number $\lambda \geq 0$.

(a) Is the robust problem (3) convex? You should be able to justify your answer based on the expression (3), without having to do any manipulations,

Solution:

The robust problem in (3) is convex, since the objective function is the pointwise maximum (over $\vec{u}$) of the convex functions, $x \to \|\vec{v} + \vec{u} - x\vec{1}\|_2^2$.

(b) Show that problem (3) can be expressed as

$$\min_{x \in \mathbb{R}} \sum_{i=1}^{n} \left(|v_i - x| + \lambda\right)^2.$$

Solution:

For a given vector $\vec{z} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq \lambda} \|\vec{z} + \vec{u}\|_2^2 &= \max_{u_i \,:\, |u_i| \leq \lambda, 1 \leq i \leq m} \sum_{i=1}^{n} (z_i + u_i)^2 = \sum_{i=1}^{n} \max_{u_i \,:\, |u_i| \leq \lambda} (z_i + u_i)^2 \\
&= \sum_{i=1}^{n} (|z_i| + \lambda)^2,
\end{aligned}$$

the last line resulting from

$$\forall\, \eta, \quad |\eta| \leq \lambda \Longrightarrow |z_i + \eta| \leq |z_i| + \lambda,$$

with the upper bound attained with $\eta = \lambda \mathrm{sign}(z_i)$.

(c) Express problem (2) as an LP. State precisely the variables and the constraints if any.

Solution:

$$\begin{aligned}
\min_x \|\vec{v} - x\vec{1}\|_1 &= \min_x \sum_{i=1}^n |v_i - x| \\
&= \min_{x,\vec{t}} \sum_{i=1}^n t_i : t_i \geq |v_i - x|, \quad \forall i \\
&= \min_{x,\vec{t}} \sum_{i=1}^n t_i : t_i \geq v_i - x, \quad t_i \geq -v_i + x, \quad \forall i.
\end{aligned}$$

(d) Express problem (3) as a QP. State precisely the variables and the constraints, if any.

Solution: Introducing slack variables as in the preceding part of the question, a QP formulation is

$$\min_{x,\vec{t}} \sum_{i=1}^n (t_i + \lambda)^2 \ : \ t_i \geq (v_i - x), \ t_i \geq -v_i + x, \ \ i = 1, \ldots, n.$$

(e) Show that when λ is large the solution set of the problem in (3) approaches that of the median problem (2).

Solution:

The objective function takes the form

$$\sum_{i=1}^n (|v_i - x| + \lambda)^2 = n\lambda^2 + 2\lambda\|\vec{v} - x\vec{1}\|_1 + \|\vec{v} - x\vec{1}\|_2^2,$$

The corresponding optimization problem has the same minimizers as the problem

$$\min_x \|\vec{v} - x\vec{1}\|_1 + \frac{1}{2\lambda}\|\vec{v} - x\vec{1}\|_2^2,$$

When λ is large, the minimizer will tend to minimize the first term only, which implies the desired result.

(f) It is often said that the median is a more robust notion of "middle" value of a finite set of real numbers than the average, when noise is present in the observations. Based on the previous part of this question, justify this statement.

Solution:

The median problem can be interpreted as a robust version of the average problem, when the uncertainty is large.

2. A matrix problem with strong duality

This problem discusses a convex optimization problem arising from the perturbation analysis of dynamical systems.

Consider the problem

$$p^* \doteq \min_{\Delta} \vec{c}^\top (A + \Delta)^{-1}\vec{b} \ : \ \|\Delta\| \leq 1,$$

where $A \in \mathbb{R}^{n \times n}$, with smallest singular value $\sigma_{\min}(A)$ strictly greater than one, and $\vec{b}, \vec{c} \in \mathbb{R}^n$ with $\vec{b}, \vec{c} \neq 0$. Here, $\|\cdot\|$ stands for the largest singular value norm, i.e. the spectral norm. This problem arises in the study of equilibrium states of a dynamical system subject to perturbations.

(a) Show that the objective function is well-defined everywhere on the feasible set.

Hint: You can show that a square matrix is invertible if it has no singular values equal to 0.

Solution:

If $\vec{y} = (A + \Delta)\vec{x}$ for $\vec{x}$ with $\|\vec{x}\|_2 = 1$, and Δ with $\|\Delta\| \leq 1$, then

$$\|\vec{y}\|_2 \geq \|A\vec{x}\|_2 - \|\Delta\vec{x}\|_2 \geq \|A\vec{x}\|_2 - \|\vec{x}\|_2 > 0,$$

due to $\sigma_{\min}(A) > 1$. This means that the nullspace of $A + \Delta$ is $\{0\}$, i.e. that $A + \Delta$ is invertible. The objective function is therefore well-defined on the feasible set.

(b) Is the problem, as stated, convex? Give a proof or a counter-example.

Solution:

The problem, as stated, is not convex in general. Indeed, for $n = 1$ the function reduces to

$$f(\Delta) = \frac{cb}{A + \Delta}, \quad |\Delta| \leq 1.$$

which can be concave, for instance, for $c = -b = 1$, $A = 2$.

(c) Show that the problem can be expressed as

$$\min_{\vec{x}} \; \vec{c}^T \vec{x} \; : \; \|A\vec{x} - \vec{b}\|_2^2 \leq \|\vec{x}\|_2^2.$$

Solution:

We will show that every feasible point in the new problem corresponds to a feasible point in the original problem with the same objective value, and vice versa, thus proving the equivalence of the two problems.

Given Δ satisfying $\|\Delta\| \leq 1$, let $\vec{x} = (A + \Delta)^{-1}\vec{b}$. We have $(A + \Delta)\vec{x} = \vec{b}$, so that $\vec{y} := b - A\vec{x} = \Delta\vec{x}$ satisfies $\|\vec{y}\|_2^2 = \|\Delta\vec{x}\|_2^2 \leq \|\vec{x}\|_2^2$. Thus, $\vec{x}$ is feasible in the new problem, with objective value $\vec{c}^\top\vec{x} = \vec{c}^\top(A + \Delta)^{-1}\vec{b}$.

Given $\vec{x}$ that is feasible for the new problem, we have $\|A\vec{x} - \vec{b}\|_2^2 \leq \|\vec{x}\|_2^2$. Let

$$\Delta := (\vec{b} - A\vec{x})(\vec{x}/\|\vec{x}\|_2^2)^T.$$

The SVD of Δ is

$$\Delta = \left(\frac{b - A\vec{x}}{\|\vec{b} - A\vec{x}\|_2}\right) \frac{\|\vec{b} - A\vec{x}\|_2}{\|\vec{x}\|_2} \left(\frac{\vec{x}}{\|\vec{x}\|_2}\right)^T,$$

with $\dfrac{\|\vec{b} - A\vec{x}\|_2}{\|\vec{x}\|_2} \leq 1$, so $\|\Delta\| \leq 1$. Thus Δ is feasible in the original problem. Further, we have $\Delta\vec{x} = \vec{b} - A\vec{x}$ so $\vec{c}^T\vec{x} = \vec{c}^T(A + \Delta)^{-1}\vec{b}$, and so the value of the objective at Δ in the old problem is the same as the objective value $\vec{c}^T\vec{x}$.

(d) Let $K := A^TA - I$. Since $\sigma_{\min}(A) > 1$, we know that K is invertible. Prove that

$$AK^{-1}A^T - I = (AA^T - I)^{-1}.$$

Solution:

$$\begin{aligned}
(AK^{-1}A^T - I)(AA^T - I) &= AK^{-1}A^TAA^T - AK^{-1}A^T - AA^T + I \\
&= A(K^{-1}A^TA - K^{-1} - I)A^T + I \\
&= A(K^{-1}(A^TA - I) - I)A^T + I \\
&= A(K^{-1}K - I)A^T + I \\
&= I
\end{aligned}$$

So, $AK^{-1}A^T - I = (AA^T - I)^{-1}$.

(This is a version of the so-called *matrix inversion lemma*.)

(e) Show that the feasible set of the formulation in (c) is an ellipsoid, expressing it in terms of the matrix $K := A^T A - I$, the vector $\vec{x}_0 := K^{-1} A^T \vec{b}$, and the scalar $\gamma := \vec{x}_0^\top K \vec{x}_0 - \vec{b}^T \vec{b}$. Explain why the above problem (which we called the new problem) is convex.

Solution:

The feasible set can be expressed as $h(\vec{x}) \leq 0$, with

$$h(\vec{x}) := \|A\vec{x} - \vec{b}\|_2^2 - \|\vec{x}\|_2^2 \tag{4}$$
$$= \vec{x}^\top K \vec{x} - 2\vec{x}^\top A^\top \vec{b} + \vec{b}^\top \vec{b} \tag{5}$$
$$= (\vec{x} - \vec{x}_0)^\top K (\vec{x} - \vec{x}_0) - \gamma, \tag{6}$$

where $K = A^\top A - I$, $\vec{x}_0 = K^{-1} A^\top \vec{b}$, $\gamma = \vec{x}_0^\top K \vec{x}_0 - \vec{b}^\top \vec{b}$.

$K \in \mathbb{S}^n_{++}$ because $\sigma_{\min}(A) > 1$. We also claim that $\gamma > 0$, which follows from

$$\begin{aligned}\gamma &= \vec{b}^T A K^{-1} A^T \vec{b} - \vec{b}^T \vec{b} \\ &= \vec{b}^T (A K^{-1} A^T - I) \vec{b} \\ &= \vec{b}^T (A A^T - I)^{-1} \vec{b},\end{aligned}$$

where we have used the result of preceding part of this question in the last step, and the observation that $(AA^T - I)^{-1} \in \mathbb{S}^n_{++}$, which also follows from $\sigma_{\min}(A) > 1$.

This establishes that the feasible set of the new problem is an ellipsoid.

The problem is one of minimizing a linear objective over an ellipsoid, so it is convex.

(f) Form a Lagrange dual to the problem. Does strong duality hold?

Solution: The dual function can be written in terms of the Lagrangian, after some algebra, as

$$g(\lambda) = \min_{\vec{x}} \; c^\top \vec{x} + \lambda((\vec{x} - \vec{x}_0)^\top K (\vec{x} - \vec{x}_0) - \gamma)$$

Since K is positive definite this equals $-\infty$ for $\lambda < 0$, and since $\vec{c} \neq \vec{0}$ it also equals $-\infty$ for $\lambda = 0$. If $\lambda > 0$, we can solve for an optimal $\vec{x}$, to find the unique optimal point

$$\vec{x}(\lambda) = \vec{x}_0 - \frac{1}{2\lambda} K^{-1} \vec{c},$$

and optimal value

$$g(\lambda) = \vec{c}^\top \vec{x}_0 - \frac{1}{4\lambda} \vec{c}^\top K^{-1} \vec{c} - \lambda \gamma.$$

The dual problem is

$$\max_{\lambda \geq 0} \; g(\lambda),$$

where we have $g(0) = -\infty$. Strong duality holds, due to the Slater condition applied to the primal, which is in turn due to the fact that the interior of the ellipsoid is not empty, since $\gamma > 0$.

(g) Show that the optimal value can be written

$$p^* = \vec{c}^\top (A^\top A - I)^{-1} A^\top \vec{b} - \|(AA^\top - I)^{-1/2} \vec{b}\|_2 \cdot \|(A^\top A - I)^{-1/2} \vec{c}\|_2.$$

Solution:

We can solve for the optimal $\lambda > 0$ in the dual problem, and we obtain this optimal value as $\vec{c}^\top \vec{x}_0 - \sqrt{\gamma \cdot \vec{c}^\top K^{-1} \vec{c}}$. Applying the matrix inversion lemma, whereby

$$AK^{-1}A^\top - I = (AA^\top - I)^{-1},$$

we obtain the desired formula.