Department of Mathematics, University of Utah
Introduction to Optimization
MATH 5770/6640, ME EN 6025 – Section 001 – Fall 2021
Homework 6
Convex optimization

Due December 7, 2021

Submit your homework assignment as a scanned copy **ON CANVAS**, to the `Homework 6` assignment.

Text: *Introduction to Nonlinear Optimization*, Amir Beck,

Exercises: # 8.2,
8.3,
8.4 (i), (ii), (iii), and only part (a) for each
11.2,
11.6 (iii) only,
P1 (**6000-level students only**)

8.2. Let $C = B[\boldsymbol{x}_0, r]$, where $\boldsymbol{x}_0 \in \mathbb{R}^n$ and $r > 0$ are given. Find a formula for the orthgonal projection operator P_C.

Solution: Note that we trivially have,

$$P_C \boldsymbol{x} = \boldsymbol{x}, \qquad\qquad \text{if } \boldsymbol{x} \in C.$$

Therefore, we assume in what follows that $\boldsymbol{x} \notin C$. In this case, we have,

$$P_C \boldsymbol{x} \coloneqq \underset{\boldsymbol{y}\in C}{\operatorname{argmin}} \|\boldsymbol{x} - \boldsymbol{y}\|_2 = \underset{\boldsymbol{y}\in C}{\operatorname{argmin}} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2,$$

However, we can restrict the minimization above to the boundary of C: a necessary condition for $\boldsymbol{y}^*$ to solve the above problem on the interior of C is that $\boldsymbol{y}^*$ must be a stationary point for the objective function $f(\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_2^2$. But f has a single stationary point at $\boldsymbol{y} = \boldsymbol{x}$, which we assume is not in C, so $\boldsymbol{y}^* = \boldsymbol{x}$ is impossible. Thus the minimizer cannot lie on the interior and must lie on the boundary. I.e., we can restrict our minimization to vectors $\boldsymbol{y}$ given by $\boldsymbol{y} = \boldsymbol{x}_0 + r\boldsymbol{a}$ for arbitrary unit vector $\boldsymbol{a} \in \mathbb{R}^n$. Therefore,

$$\begin{aligned}
\min_{\boldsymbol{y}\in\partial C} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 &= \min_{\|\boldsymbol{a}\|_2=1} \|\boldsymbol{x} - \boldsymbol{x}_0 - r\boldsymbol{a}\|_2^2, \\
&= \min_{\|\boldsymbol{a}\|_2=1} \left\langle \boldsymbol{x} - \boldsymbol{x}_0 - r\boldsymbol{a}, \boldsymbol{x} - \boldsymbol{x}_0 - r\boldsymbol{a} \right\rangle \\
&= \min_{\|\boldsymbol{a}\|_2=1} \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{x}_0\|_2^2 + r^2\|\boldsymbol{a}\|_2^2 - 2\left\langle \boldsymbol{x}, \boldsymbol{x}_0 \right\rangle - 2r\left\langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{x}_0 \right\rangle \\
&= \min_{\|\boldsymbol{a}\|_2=1} -2r\left\langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{x}_0 \right\rangle \\
&= \max_{\|\boldsymbol{a}\|_2=1} \left\langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{x}_0 \right\rangle.
\end{aligned}$$

By the Cauchy-Schwarz inequality, the solution of the above maximization is,

$$\underset{\|\boldsymbol{a}\|_2=1}{\operatorname{argmax}} = \frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|_2}.$$

Therefore, we have

$$P_C\boldsymbol{x} = \begin{cases} \boldsymbol{x}, & \boldsymbol{x} \in C \\ \boldsymbol{x}_0 + r\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|_2}, & \boldsymbol{x} \notin C \end{cases}$$

8.3. Let f be a strictly convex function over $\mathbb{R}^m$ and let g be a convex function over $\mathbb{R}^n$. Define the function,

$$h(\boldsymbol{x}) = f(\boldsymbol{A}\boldsymbol{x}) + g(\boldsymbol{x}),$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Assume that $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ are optimal solutions of the unconstrained problem of minimizing h. Show that $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{y}^*$.

Solution: The function h is a sum of convex functions, and is hence convex itself. Therefore, the set of minimizers is itself convex. Thus, the entire line segment $[\boldsymbol{x}^*, \boldsymbol{y}^*]$ is a minimizer of h. (Recall that $[\boldsymbol{x}^*, \boldsymbol{y}^*]$ means the closed line segment connecting $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ in $\mathbb{R}^n$.) Therefore, for any $\lambda \in [0, 1]$, we have,

$$\lambda\boldsymbol{x}^* + (1-\lambda)\boldsymbol{y}^* \in \underset{\boldsymbol{x}\in\mathbb{R}^n}{\operatorname{argmin}}\, h(\boldsymbol{x}),$$

and we also have $h(\boldsymbol{x}^*) = h(\boldsymbol{y}^*)$. Now assume that $\boldsymbol{A}\boldsymbol{x}^* \neq \boldsymbol{A}\boldsymbol{y}^*$; we will prove the desired result by contradiction. Using convexity of g and strict convexity of f, we have for any $\lambda \in (0, 1)$,

$$\begin{aligned} h(\lambda\boldsymbol{x}^* + (1-\lambda)\boldsymbol{y}^*) &= f(\lambda\boldsymbol{A}\boldsymbol{x}^* + (1-\lambda)\boldsymbol{A}\boldsymbol{y}^*) + g(\lambda\boldsymbol{x}^* + (1-\lambda)\boldsymbol{y}^*) \\ &\leq f(\lambda\boldsymbol{A}\boldsymbol{x}^* + (1-\lambda)\boldsymbol{A}\boldsymbol{y}^*) + \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}) \\ &\overset{(*)}{<} \lambda f(\boldsymbol{A}\boldsymbol{x}^*) + (1-\lambda)f(\boldsymbol{A}\boldsymbol{y}^*) + \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}) \\ &= \lambda h(\boldsymbol{x}^*) + (1-\lambda)h(\boldsymbol{y}^*) \\ &= h(\boldsymbol{x}^*). \end{aligned}$$

where the strict inequality $(*)$ is true since f is strictly convex and we assume $\boldsymbol{A}\boldsymbol{x}^* \neq \boldsymbol{A}\boldsymbol{y}^*$. We have shown that h evaluated at $\lambda\boldsymbol{x}^* + (1-\lambda)\boldsymbol{y}^*$ is strictly smaller than h at $\boldsymbol{x}^*$, which contradicts the optimality of $\boldsymbol{x}^*$. Therefore, our assumption $\boldsymbol{A}\boldsymbol{x}^* \neq \boldsymbol{A}\boldsymbol{y}^*$ must be incorrect, so that $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{y}^*$.

8.4. For each of the following optimization problems, (a) show that it is convex. (Note: there are parts (b) and (c) of this problem that were not assigned.)

(i)

$$\begin{aligned} \min \quad & x_1^2 + 2x_1x_2 + 2x_2^2 + x_3^2 + 3x_1 - 4x_2 \\ \text{s.t.} \quad & \sqrt{2x_1^2 + x_1x_2 + 4x_2^2 + 4} + \frac{(x_1 - x_2 + x_3 + 1)^2}{x_1 + x_2} \leq 6 \\ & x_1, x_2, x_3 \geq 1. \end{aligned}$$

(ii)

$$\begin{aligned} \max \quad & x_1 + x_2 + x_3 + x_4 \\ \text{s.t.} \quad & (x_1 - x_2)^2 + (x_3 + 2x_4)^4 \leq 5 \\ & x_1 + 2x_2 + 3x_3 + 4x_4 \leq 6 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

(iii)

$$\begin{aligned} \min \quad & 5x_1^2 + 4x_2^2 + 7x_3^2 + 4x_1x_2 + 2x_2x_3 + |x_1 - x_2| \\ \text{s.t.} \quad & \frac{x_1^2 + x_2^2}{x_3} + (x_1^2 + x_2^2 + 1)^4 \leq 10 \\ & x_3 \geq 10 \end{aligned}$$

Solution: In all the following solutions, we have shown the problem is convex if we demonstrate that the objective function is convex, and the constaint set is convex.

(i) The objective function can be written as,

$$x_1^2 + 2x_1x_2 + 2x_2^2 + x_3^2 + 3x_1 - 4x_2 = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} + (3x_1 - 4x_2),$$

where

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \succ \boldsymbol{0},$$

and therefore the objective function is a conic combination of convex functions (all affine functions are convex) and hence the objective function is convex.

The full constraint set is an intersection of 4 convex sets: if we show that each set is convex, then their intersection is also convex. The last three constraints, $x_1, x_2, x_3 \geq 1$ are convex sets (e.g., they are level sets of convex, affine functions). The first constraint can be written as,

$$\sqrt{2x_1^2 + x_1x_2 + 4x_2^2 + 4} + \frac{(x_1 - x_2 + x_3 + 1)^2}{x_1 + x_2} = 2\sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 1} + \frac{\|\boldsymbol{B}\boldsymbol{x} + \boldsymbol{c}\|_2^2}{\boldsymbol{w}^T \boldsymbol{x} + 0},$$

where

$$\boldsymbol{Q} = \begin{pmatrix} \frac{1}{2} & \frac{1}{8} & 0 \\ \frac{1}{8} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \succeq \boldsymbol{0}, \qquad \boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 \end{pmatrix}, \qquad \boldsymbol{w} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, this constraint is the conic sum of a function of the form $\sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 1}$ with $\boldsymbol{Q} \succeq \boldsymbol{0}$ (which is convex by homework #5), and a quadratic-over-linear function (which is convex as shown in Lecture 14). Therefore, this constraint is also convex.

(ii) The objective function is affine (actually linear), and therefore is convex. The last 4 constraints $x_1, x_2, x_3, x_4 \geq 0$ are all convex being level sets of affine functions. The penultimate constraint $x_1 + 2x_2 + 3x_3 + 4x_4 \le 6$ is the level set of an affine function, and is hence convex. The first constraint is the level set of the function,

$$f(\boldsymbol{x}) = (x_1 - x_2)^2 + (x_3 + 2x_4)^4,$$

so we need only show that f is convex. The function f is a conic combination of two functions, so we focus on showing that these two functions are individually convex, which would prove the result. The function,

$$g(\boldsymbol{x}) = (x_1 - x_2)^2 = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}, \qquad \boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \succeq \boldsymbol{0},$$

and is hence convex. The other function can be written as,

$$h(\boldsymbol{x}) = (x_3 + 2x_4)^4 = (\boldsymbol{B}\boldsymbol{x})^4, \qquad \boldsymbol{B} = \begin{pmatrix} 0 & 0 & 1 & 2 \end{pmatrix},$$

which is a composition of the strictly convex function $t \mapsto t^4$ with an affine function $\boldsymbol{x} \mapsto \boldsymbol{B}\boldsymbol{x}$. Hence, this composed function h is convex (cf. Theorem 7.17).

(iii) The objective function can be written as,

$$5x_1^2 + 4x_2^2 + 7x_3^2 + 4x_1x_2 + 2x_2x_3 + |x_1 - x_2| = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + |\boldsymbol{a}^T \boldsymbol{x}|,$$

where,

$$\boldsymbol{Q} = \begin{pmatrix} 5 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 1 & 7 \end{pmatrix} \succ \boldsymbol{0}, \qquad \boldsymbol{a}^T = (1, -1, 0).$$

Hence, the objective function is a conic combination of a positive-definite quadratic function ($\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}$) and a second function that is also convex since it's the composition of a convex function with an affine function. Thus, the objective function is convex.

The last constraint, $x_3 \ge 10$, is a convex set since it's the level set of an affine function. The first constraint is the level set of the function,

$$f(\boldsymbol{x}) = \frac{x_1^2 + x_2^2}{x_3} + (x_1^2 + x_2^2 + 1)^4,$$

so we seek to show that this function is convex. Write this function as a conic combination of two functions,

$$f(\boldsymbol{x}) = \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2^2}{\boldsymbol{w}^T \boldsymbol{x} + 0} + \left(\boldsymbol{x}^T \boldsymbol{R} \boldsymbol{x} + 1\right)^4,$$

where

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \succeq \boldsymbol{0}$$

The first function is a quadratic-over-linear function, and hence is convex. The second function is the composition of the function $t \mapsto t^4$, which is monotone increasing and convex for $t \ge 0$, with the quadratic function $\boldsymbol{x}^T \boldsymbol{R} \boldsymbol{x} + 1$, which is convex since $\boldsymbol{R} \succeq \boldsymbol{0}$. Therefore, this second function is also convex. Thus, f is the conic combination of two convex functions, and is thus convex.

11.2. Consider the optimization problem

$$\text{(P)} \quad \min \left\{ \boldsymbol{a}^T \boldsymbol{x} : \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c \le 0 \right\},$$

where $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is positive definite, $\boldsymbol{a}(\neq \boldsymbol{0})$, $\boldsymbol{b} \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(i) For which values of $\boldsymbol{Q}$, $\boldsymbol{b}$, c is the problem infeasible?
(ii) For which values of $\boldsymbol{Q}$, $\boldsymbol{b}$, c are the KKT conditions necessary?
(iii) For which values of $\boldsymbol{Q}$, $\boldsymbol{b}$, c are the KKT conditions sufficient?
(iv) Under the condition of part (ii), find the optimal solution of (P) using the KKT conditions.

Solution:

(i) Since $\boldsymbol{Q} \succ \boldsymbol{0}$, the feasible set is a compact set, and the objective function is continuous. So the problem is feasible if and only if the feasible set is non-empty, i.e., there must exist at least one $\boldsymbol{x}$ such that $\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c \le 0$. This happens exactly when the minimum value of the function g defined as,

$$g(\boldsymbol{x}) \coloneqq \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c,$$

is non-negative. Since this is a quadratic function (with a positive-definite Hessian $2\boldsymbol{Q}$), the unique minimum of this function is the unique solution to the first-order conditions $\nabla g(\boldsymbol{x}) = \boldsymbol{0}$. This minimum is given by,

$$-\boldsymbol{Q}^{-1}\boldsymbol{b} = \tilde{\boldsymbol{x}} \coloneqq \underset{\boldsymbol{x}\in\mathbb{R}^n}{\operatorname{argmin}}\, g(\boldsymbol{x}).$$

We then directly compute,

$$g(\tilde{\boldsymbol{x}}) = c - \boldsymbol{b}^T \boldsymbol{Q}^{-1} \boldsymbol{b}.$$

Therefore, problem (P) is feasible if and only if $g(\widetilde{\boldsymbol{x}}) \le 0$, i.e., if and only if

$$c \le \boldsymbol{b}^T \boldsymbol{Q}^{-1} \boldsymbol{b}$$

(ii) First we note that the solution to this problem lies on the *boundary* of the feasible set: if $\boldsymbol{x}^*$ is a local minimum that lies in the interior, then there is some $\epsilon > 0$ such that $B(\boldsymbol{x}^*, \epsilon)$ also lies in the interior, and therefore for any $t < \epsilon$, we have,

$$\begin{aligned} \boldsymbol{x}^* - \frac{t}{2}\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_2} &\text{ is feasible} \\ f\left(\boldsymbol{x}^* - \frac{t}{2}\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_2}\right) &= f(\boldsymbol{x}^*) - \frac{t\|\boldsymbol{a}\|_2}{2} < f(\boldsymbol{x}^*), \end{aligned}$$

where $f(\boldsymbol{x}) \coloneqq \boldsymbol{a}^T\boldsymbol{x}$. This shows that $\boldsymbol{x}^*$ cannot be a local minimum if it lies in the interior of the feasible set, and thus local minima must lie on the boundary.

The KKT conditions are necessary when local minima of the problem are regular points. Regularity in this case demands that $\nabla g(\boldsymbol{x})$ not be a linearly dependent vector, which occurs only when $\nabla g(\boldsymbol{x}) = \boldsymbol{0}$. In part (i) we already showed that there is a single point at which this happens, $\boldsymbol{x} = -\boldsymbol{Q}^{-1}\boldsymbol{b}$, and thus this is the only irregular point. Does this irregular point lie on the boundary of the feasible set? At this irregular point, we have

$$g(\boldsymbol{x}) = c - \boldsymbol{b}^T \boldsymbol{Q}^{-1} \boldsymbol{b}.$$

If this evaluates to exactly 0, then $\boldsymbol{x}$ is on the boundary of the feasible set, but if this evaluates to any (strictly) negative number, then $\boldsymbol{x}$ lies in the interior. In the latter case, although $\boldsymbol{x} = -\boldsymbol{Q}^{-1}\boldsymbol{b}$ is always irregular, it cannot be a local minimum. Therefore, the KKT conditions are necessary when

$$c < \boldsymbol{b}^T \boldsymbol{Q}^{-1} \boldsymbol{b}$$

(iii) The function $f(\boldsymbol{x}) \coloneqq \boldsymbol{a}^T\boldsymbol{x}$ is convex, and $g(\boldsymbol{x})$ is (strictly) convex since $\boldsymbol{Q}$ is (strictly) positive-definite. Therefore this is a convex optimization problem and the KKT conditions are always sufficient (assuming the problem is feasible).

(iv) Defining $f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x}$, and with g the constraint function defined above, we wish to use the KKT conditions on the problem,

$$\min f(\boldsymbol{x}) \text{ subject to } g(\boldsymbol{x}) \leq 0,$$

where we assume as determined in part (ii) that $\boldsymbol{b}^T\boldsymbol{Q}^{-1}\boldsymbol{b} - c > 0$. The KKT conditions in this case are,

$$\begin{aligned} \nabla f + \lambda \nabla g = \boldsymbol{a} + 2\lambda(\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}) &= \boldsymbol{0}, \\ \lambda g(\boldsymbol{x}) = \lambda(\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + 2\boldsymbol{b}^T\boldsymbol{x} + c) &= 0, \\ \lambda &\geq 0, \\ g(\boldsymbol{x}) &\leq 0. \end{aligned}$$

First we note a simplification: if $\lambda = 0$, then the first condition becomes $\boldsymbol{a} = \boldsymbol{0}$, which violates the assumption in the problem statement. Therefore, we must have $\lambda > 0$. This then implies through the second condition that $g(\boldsymbol{x}) = 0$, which makes the fourth condition unnecessary. Therefore, a simplification of the above is:

$$\begin{aligned} \boldsymbol{a} + 2\lambda(\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}) &= \boldsymbol{0}, \\ \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + 2\boldsymbol{b}^T\boldsymbol{x} + c &= 0, \\ \lambda &> 0. \end{aligned}$$

We now simplify this problem even further through a variable transformation: Since $\boldsymbol{Q} \succ \boldsymbol{0}$, then let

$$\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{L}^T, \qquad \boldsymbol{y} \coloneqq \boldsymbol{L}^T\boldsymbol{x},$$

where $\boldsymbol{L}$ can be either a Cholesky factor of $\boldsymbol{Q}$ or the matrix square root of $\boldsymbol{Q}$ computed through its eigenvalue decomposition. (In the latter case $\boldsymbol{L}$ is symmetric, but we do not make this assumption in what follows.) We can then rewrite our KKT conditions in terms of the new variable $\boldsymbol{y}$:

$$\widetilde{\boldsymbol{a}} + 2\lambda(\boldsymbol{y} + \widetilde{\boldsymbol{b}}) = \boldsymbol{0}, \tag{1a}$$

$$\boldsymbol{y}^T\boldsymbol{y} + 2\widetilde{\boldsymbol{b}}^T\boldsymbol{y} + c = 0, \tag{1b}$$

$$\lambda > 0, \tag{1c}$$

where we have defined new vectors $\widetilde{\boldsymbol{a}}$ and $\widetilde{\boldsymbol{b}}$ as,

$$\widetilde{\boldsymbol{a}} \coloneqq \boldsymbol{L}^{-1}\boldsymbol{a}, \qquad \widetilde{\boldsymbol{b}} \coloneqq \boldsymbol{L}^{-1}\boldsymbol{b}.$$

We will first solve (1) for $\boldsymbol{y}$ and then compute $\boldsymbol{x} = \boldsymbol{L}^T \boldsymbol{x}$. Note that (1a) implies,

$$\boldsymbol{y} + \widetilde{\boldsymbol{b}} = -\frac{\widetilde{\boldsymbol{a}}}{2\lambda}. \tag{2}$$

We now complete the square on (1b) and use (2):

$$\begin{aligned}
\boldsymbol{y}^T\boldsymbol{y} + 2\widetilde{\boldsymbol{b}}^T\boldsymbol{y} + \widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - \widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} + c = 0& \\
\left(\boldsymbol{y} + \widetilde{\boldsymbol{b}}\right)^T \left(\boldsymbol{y} + \widetilde{\boldsymbol{b}}\right) &= \widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - c \\
\left(-\frac{\widetilde{\boldsymbol{a}}}{2\lambda}\right)^T \left(-\frac{\widetilde{\boldsymbol{a}}}{2\lambda}\right) &= \widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - c \\
\frac{\|\widetilde{\boldsymbol{a}}\|^2}{4\lambda^2} &= \widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - c,
\end{aligned}$$

and noting $\lambda > 0$ and solving for 2λ yields,

$$2\lambda = \frac{\|\widetilde{\boldsymbol{a}}\|}{\sqrt{\widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - c}},$$

where we note that the factor under the square root is strictly positive owing to the necessary conditions in part (ii). Using this formula for λ in (2) allows us to compute $\boldsymbol{y}$:

$$\boldsymbol{y} = -\widetilde{\boldsymbol{b}} - \frac{\widetilde{\boldsymbol{a}}}{\|\widetilde{\boldsymbol{a}}\|}\sqrt{\widetilde{\boldsymbol{b}}^T\widetilde{\boldsymbol{b}} - c}$$

Using $\boldsymbol{y} = \boldsymbol{L}^T\boldsymbol{x}$ to solve for $\boldsymbol{x}$ yields the single KKT point,

$$\boldsymbol{x} = -\boldsymbol{Q}^{-1}\left(\boldsymbol{b} + \boldsymbol{a}\sqrt{\frac{\boldsymbol{b}^T\boldsymbol{Q}^{-1}\boldsymbol{b} - c}{\boldsymbol{a}^T\boldsymbol{Q}^{-1}\boldsymbol{a}}}\right).$$

Note that since this problem is feasible (by part (i)) and that the KKT conditions are necessary (by part (ii)) that this single KKT point must in fact be the global minimum solution to (P).

11.6(iii). Use the KKT conditions in order to find an optimal solution of the following problem:

$$\begin{aligned}
&\min\ 2x_1 + x_2 \\
&\text{s.t. } 4x_1^2 + x_2^2 - 2 \le 0 \\
&\qquad 4x_1 + x_2 + 3 \le 0.
\end{aligned}$$

Solution: Define,

$$f(\boldsymbol{x}) = 2x_1 + x_2, \qquad g_1(\boldsymbol{x}) = 4x_1^2 + x_2^2 - 2, \qquad g_2(\boldsymbol{x}) = 4x_1 + x_2 + 3,$$

and note that f, g_1, and g_2 are all convex. Therefore, the KKT conditions are sufficient to find a global minimum. First we compute,

$$\nabla f = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \nabla g_1(\boldsymbol{x}) = 2\begin{pmatrix} 4x_1 \\ x_2 \end{pmatrix}, \qquad \nabla g_2(\boldsymbol{x}) = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

The KKT conditions then read,

$$\nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = \left(\begin{array}{c} 2 \\ 1 \end{array} \right) + 2\lambda_1 \left(\begin{array}{c} 4x_1 \\ x_2 \end{array} \right) + \lambda_2 \left(\begin{array}{c} 4 \\ 1 \end{array} \right) = \mathbf{0}, \tag{3a}$$

$$\lambda_1 g_1 = \lambda_1 \left(4x_1^2 + x_2^2 - 2\right) = 0 \tag{3b}$$

$$\lambda_2 g_2 = \lambda_2 \left(4x_1 + x_2 + 3\right) = 0 \tag{3c}$$

$$\lambda_1, \lambda_2 \geq 0, \tag{3d}$$

$$g_1(\boldsymbol{x}), g_2(\boldsymbol{x}) \leq 0. \tag{3e}$$

First we note that $\lambda_1, \lambda_2 = 0$ is not possible since that violates (3a). If we consider $\lambda_1 = 0$ and $\lambda_2 \neq 0$, then (3a) reads,

$$\left(\begin{array}{c} 2 \\ 1 \end{array} \right) + \lambda_2 \left(\begin{array}{c} 4 \\ 1 \end{array} \right) = \mathbf{0},$$

which is not possible since the vectors $(2,1)^T$ and $(4,1)^T$ are linearly independent.
If we allow $\lambda_1, \lambda_2 \neq 0$, then (3b) and (3c) imply,

$$4x_1^2 + x_2^2 - 2 = 0, \qquad\qquad 4x_1 + x_2 + 3 = 0.$$

The second (linear) expression shows that $x_2 = -3 - 4x_1$, and substituting this into the quadratic condition yields, after simplification,

$$20x_1^2 + 24x_1 + 7 = 0,$$

whose two solutions are,

$$x_1 = -\frac{1}{2}, -\frac{7}{10}.$$

Now we have,

$$x_1 = -\frac{1}{2} \overset{\text{(3c)}}{\Longrightarrow} x_2 = -1 \overset{\text{(3a)}}{\Longrightarrow} (\lambda_1, \lambda_2) = \left(\frac{1}{2}, 0 \right),$$

but this violates our assumption that $\lambda_2 = 0$, so $x_1 = -\frac{1}{2}$ is not possible. On the other hand,

$$x_1 = -\frac{7}{10} \overset{\text{(3c)}}{\Longrightarrow} x_2 = -\frac{1}{5} \overset{\text{(3a)}}{\Longrightarrow} (\lambda_1, \lambda_2) = \left(-\frac{1}{2}, -\frac{6}{5} \right),$$

which violates (3d), and so $x_1 = -\frac{7}{10}$ is also not possible. Therefore, $\lambda_1, \lambda_2 \neq 0$ also yields no KKT points.
The final possibility is $\lambda_2 = 0$, $\lambda_1 \neq 0$. In this case (3a) implies,

$$\left(\begin{array}{c} 2 \\ 1 \end{array} \right) + 2\lambda_1 \left(\begin{array}{c} 4x_1 \\ x_2 \end{array} \right) = \mathbf{0},$$

which can only happen if $(2,1)^T$ and $(4x_1, x_2)^T$ are linearly dependent, i.e., only if $x_1 = \frac{1}{2}x_2$.
Using this condition for x_1 along with $\lambda_1 \neq 0$ in (3b) yields,

$$4\left(\frac{1}{4}x_2^2\right) + x_2^2 - 2 = 0 \implies x_2 = \pm 1.$$

To determine if either of these options is viable, we again consider (3a), which under the $x_1 = \frac{1}{2}x_2$ condition reads,

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} + 2\lambda_1 x_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \mathbf{0} \implies \begin{pmatrix} 2 \\ 1 \end{pmatrix} [1 + 2\lambda_1 x_2] = \mathbf{0},$$

implying in particular that $\lambda_1 = -\frac{1}{2x_2}$. This condition then disallows $x_2 = +1$. Then in the case $x_2 = -1$, we have

$$(x_1, x_2, \lambda_1, \lambda_2) = \left(-\frac{1}{2}, -1, \frac{1}{2}, 0\right),$$

which indeed satisfies (3), and hence is the only KKT point. Since the KKT conditions are sufficient, then $\boldsymbol{x} = \left(-\frac{1}{2}, -1\right)^T$ is the optimal solution.

Additional problems:

P1. (**6000-level students only**) Given $n \in \mathbb{N}$, let C be the closed, convex set of $n \times n$ positive semi-definite matrices,

$$C = \left\{ \boldsymbol{A} \in \mathbb{R}^{n\times n} \;\middle|\; \boldsymbol{A} \text{ is symmetric, } \boldsymbol{A} \succeq \mathbf{0} \right\}.$$

Given any $n \times n$ symmetric matrix $\boldsymbol{A}$, let $P_C(\boldsymbol{A})$ denote the $\|\cdot\|$-projection of $\boldsymbol{A}$ onto C. Show that both of the following statements are true,

$$\begin{aligned} P_C(\boldsymbol{A}) &= \boldsymbol{U}\boldsymbol{\Lambda}_+\boldsymbol{U}^T, && \|\cdot\| = \|\cdot\|_2 \text{ (the spectral or induced matrix 2-norm)} \\ P_C(\boldsymbol{A}) &= \boldsymbol{U}\boldsymbol{\Lambda}_+\boldsymbol{U}^T, && \|\cdot\| = \|\cdot\|_F \text{ (the Frobenius, or entrywise norm)} \end{aligned}$$

where $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$ is the eigenvalue decomposition of $\boldsymbol{A}$, and $\boldsymbol{\Lambda}_+ = \max\{\boldsymbol{\Lambda}, 0\}$ with the max function applied componentwise.

Solution: First we consider the 2-norm, $\|\cdot\| = \|\cdot\|_2$. Given the symmetric matrix $\boldsymbol{A}$, let

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T = \sum_{j=1}^{n} \lambda_j \boldsymbol{u}_j \boldsymbol{u}_j^T,$$

denote the eigenvalue decomposition of $\boldsymbol{A}$, where $\{\lambda_j\}_{j=1}^n$ are the eigenvalues (also the diagonal elements of $\boldsymbol{\Lambda}$) and $\{\boldsymbol{u}\}_{j=1}^n$ are the unit-norm eigenvectors (also the columns of $\boldsymbol{U}$). We assume without loss that the eigenvalues are ordered, i.e.,

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n,$$

Note that we may assume $\lambda_n < 0$ without loss, since otherwise, $\boldsymbol{A}$ is itself a positive semidefinite matrix, and $P_C(\boldsymbol{A}) = \boldsymbol{A}$. First note the following property of the 2-norm for symmetric matrices,

$$\|\boldsymbol{A}\|_2^2 = \sup_{\|\boldsymbol{u}\|_2 = 1} \|\boldsymbol{A}\boldsymbol{u}\|_2^2 = \sup_{\|\boldsymbol{u}\|_2 = 1} R_{\boldsymbol{A}^T\boldsymbol{A}}(\boldsymbol{u}) = \max_{i=1,\ldots,n} |\lambda_i|^2 = \max\{\lambda_1, -\lambda_n\},$$

where $R_{\boldsymbol{V}}(\boldsymbol{u})$ is the Rayleigh quotient of $\boldsymbol{V}$ at vector $\boldsymbol{u}$, and we have used the fact that the eigenvalue decomposition of $\boldsymbol{A}^T\boldsymbol{A}$ is $\boldsymbol{U}\boldsymbol{\Lambda}^2\boldsymbol{U}^T$ to maximize the Rayleigh quotient.

Now let $\boldsymbol{B} \in C$ be arbitrary. Using the above property, and recalling that $\boldsymbol{v}_n$ is the eigenvector of $\boldsymbol{A}$ associated to λ_n, we have

$$\|\boldsymbol{B}-\boldsymbol{A}\|_2 \geq \sup_{\|\boldsymbol{v}\|_2=1} R_{\boldsymbol{B}-\boldsymbol{A}}(\boldsymbol{v}) \geq \boldsymbol{v}_n^T(\boldsymbol{B}-\boldsymbol{A})\boldsymbol{v}_n = \boldsymbol{v}_n^T\boldsymbol{B}\boldsymbol{v}_n^T - \lambda_n \geq -\lambda_n,$$

where the last inequality is true since $\boldsymbol{B}$ is positive semi-definite. Thus, $\|\boldsymbol{B}-\boldsymbol{A}\|_2 \geq |\lambda_n|$. But we can achieve this lower bound:

$$\boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}_+\boldsymbol{U}^T \implies \|\boldsymbol{A}-\boldsymbol{B}\|_2 = \left\|\sum_{j=1}^n \max\{0,-\lambda_j\}\boldsymbol{u}_j\boldsymbol{u}_j^T\right\|_2 = \max_{i=1,\ldots,n} \max\{0,-\lambda_j\} = -\lambda_n.$$

I.e., our choice of $\boldsymbol{B}$ achieves the desired lower bound distance for every element in C, and thus $\boldsymbol{B}$ is the projection onto C in the $\|\cdot\|_2$ norm.
Now we consider the Frobenius norm, $\|\cdot\| = \|\cdot\|_F$. Again let $\boldsymbol{B} \in C$ be arbitrary. Since the Frobenius norm is invariant with respect to unitary transforms, we have,

$$\|\boldsymbol{B}-\boldsymbol{A}\|_F^2 = \left\|\boldsymbol{U}^T\boldsymbol{B}\boldsymbol{U}-\boldsymbol{\Lambda}\right\|_F^2.$$

Note that $\boldsymbol{X} \coloneqq \boldsymbol{U}^T\boldsymbol{B}\boldsymbol{U}$ is also positive semi-definite, and hence is also an element of C. So we seek to minimize,

$$\|\boldsymbol{X}-\boldsymbol{\Lambda}\|_F^2, \qquad\qquad \boldsymbol{X} = \boldsymbol{U}^T\boldsymbol{B}\boldsymbol{U} \in C.$$

Recall that if $\boldsymbol{X}$ is positive semi-definite, then its diagonal entries must also be non-negative. Therefore,

$$\|\boldsymbol{X}-\boldsymbol{\Lambda}\|_F^2 = \sum_{j=1}^n (X_{j,j}-\lambda_j)^2 + \sum_{i\neq j} X_{i,j}^2 \geq \sum_{j=1}^n (\max\{0,-\lambda_j\})^2 + \sum_{i\neq j} X_{i,j}^2 \geq \sum_{j=1}^n (\max\{0,-\lambda_j\})^2$$

which is achieved by selecting the diagonal elements of $\boldsymbol{X}$ to equal the diagonal elements of $\boldsymbol{\Lambda}_+$, and by choosing the off diagonal elements of $\boldsymbol{X}$ to vanish. Thus, we can achieve the lower bound,

$$\|\boldsymbol{X}-\boldsymbol{\Lambda}\|_F^2 \geq \sum_{j=1}^n (\max\{0,-\lambda_j\})^2,$$

by choosing $\boldsymbol{X} = \boldsymbol{\Lambda}_+$. Hence,

$$\operatorname*{argmin}_{\boldsymbol{X}\in C} \|\boldsymbol{X}-\boldsymbol{\Lambda}\|_F^2 = \boldsymbol{\Lambda}_+,$$

implying in turn that,

$$P_C(\boldsymbol{A}) = \operatorname*{argmin}_{\boldsymbol{B}\in C} \|\boldsymbol{B}-\boldsymbol{A}\|_F^2 = \boldsymbol{U}\left(\operatorname*{argmin}_{\boldsymbol{X}\in C} \|\boldsymbol{X}-\boldsymbol{\Lambda}\|_F^2\right)\boldsymbol{U}^T = \boldsymbol{U}\boldsymbol{\Lambda}_+\boldsymbol{U}^T,$$

proving the result.