EE578B - Convex Optimization - Winter 2021

Homework 3 - Solution

Due Date: Sunday, Jan 31^{st}, 2020 at 11:59 pm

1. **Quadratic Functions**

   Consider the quadratic function

   $$f(x) = \tfrac{1}{2}x^T Q x + c^T x$$

   - **(PTS:0-2)**Rewrite $f(x)$ in the form

     $$f(x) = \tfrac{1}{2}(x - x_c)^T Q (x - x_c) + \text{CONST}$$

     Solution: Working backwards we get...

     $$\tfrac{1}{2}(x - x_c)^T Q (x - x_c) = \tfrac{1}{2}x^T Q x - x_c^T Q x + \tfrac{1}{2}x_c^T Q x_c$$

     It follows that

     $$-x_c^T Q x = c^T x \qquad \Rightarrow \qquad x_c = -Q^{-1}c$$

     (Q is symmetric.) Thus we can write

     $$\begin{aligned} f(x) &= \tfrac{1}{2}x^T Q x + c^T x \\ &= \tfrac{1}{2}(x + Q^{-1}c)^T Q (x + Q^{-1}c) - \tfrac{1}{2}c^T Q^{-1} c \end{aligned}$$

   - **(PTS:0-2)** Compute the derivative of both forms of $f(x)$ and show that they are the same.
     Solution:
     Using the original form of $f(x)$ we get

     $$\frac{\partial f}{\partial x} = x^T Q + c^T$$

     Using the second form (using the chain rule) we get

     $$(x + Q^{-1}c)^T Q = x^T Q + c^T$$

2. **Minimum Norm Problem**

   Consider the following optimization problem for finding the minimum norm solution to a linear system of equations

   $$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) = \tfrac{1}{2}|x|_2^2 = \tfrac{1}{2}x^T x \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

for $A \in \mathbb{R}^{m\times n}$ full row rank with $m < n$ and $b \in \mathbb{R}^m$. The optimality conditions for this optimization problem are given by

$$\frac{\partial f}{\partial x}^T = x = -A^T v \tag{1}$$
$$Ax = b \tag{2}$$

with dual variable $v \in \mathbb{R}^m$. Let x^*, v^* refer to x and v at optimum.

- **(PTS:0-2)** Solve for v^* in terms of b. (Hint: start by left multiplying (1) by A and substituting in $Ax = b$).
  Solution:

$$\begin{aligned} x &= -A^T v \\ Ax &= -AA^T v \\ b &= -AA^T v \\ \Rightarrow \qquad v &= -(AA^T)^{-1} b \end{aligned}$$

- **(PTS:0-2)** Solve for x^* in terms of b.
  Solution:

$$x^* = A^T(AA^T)^{-1}b$$

- **(PTS:0-2)** Let the columns of $N \in \mathbb{R}^{n\times(n-m)}$ form a basis for the nullspace of A. Compute $z_1^* \in \mathbb{R}^m$ and $z_2^* \in \mathbb{R}^{n-m}$ such that

$$x^* = \underbrace{\begin{bmatrix} A^T & N \end{bmatrix}}_{P} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix}$$

  ie. write x^* in terms of the coordinates with respect to the columns of P. Interpret z_1^* and z_2^* in terms of projections of x^* onto $\mathcal{R}(A^T)$ and $\mathcal{R}(N)$. How does z_1^* relate to v^*? Explain the value of z_2^* intuitively.
  Solution:
  From the previous homework we have that

$$\begin{bmatrix} A^T & N \end{bmatrix}^{-1} = \begin{bmatrix} (AA^T)^{-1}A \\ (N^TN)^{-1}N^T \end{bmatrix}$$

  Thus we can write

$$\begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} = \begin{bmatrix} (AA^T)^{-1}Ax^* \\ (N^TN)^{-1}N^Tx^* \end{bmatrix}$$

  $A^Tz_1^* = A^T(AA^T)^{-1}Ax^*$ is the projection of x^* onto $\mathcal{R}(A^T)$ and $Nz_2^* = N(N^TN)^{-1}N^Tx^*$ is the projection of x^* onto $\mathcal{N}(A)$. At optimum, we have that $x^* = A^Tz_1^* + Nz_2^*$, ie. z_1^* and z_2^* are each components of the coordinates of x^* with respect to the basis $[A^T \ N]$. Since $x^* = -A^Tv$, we have that $z_1^* = -v$ and $z_2^* = 0$. Intuitively, we are trying to find x with the smallest 2-norm such that $Ax = b$. We get the smallest norm x by not including any component of x in the nullspace of A, ie. $z_2^* = 0$ since setting $z_2^* \neq 0$ only increases the norm of x^* without changing the quantity $Ax^* = A(A^Tz_1^* + Nz_2^*) = AA^Tz_1^* + ANz_2^* = AA^Tz_1^*$.

- **(PTS:0-2)** Consider the above problem for $A = [1\ 1]$ and $b = 1$. Draw a picture of the optimization space labeling

$$\{x \in \mathbb{R}^2 \mid Ax = b\}, \quad \mathcal{R}(A^T), \quad x^*, \quad -A^T v^*, \quad \text{level sets of } f(x), \quad \left.\frac{\partial f}{\partial x}\right|_{x^*}$$

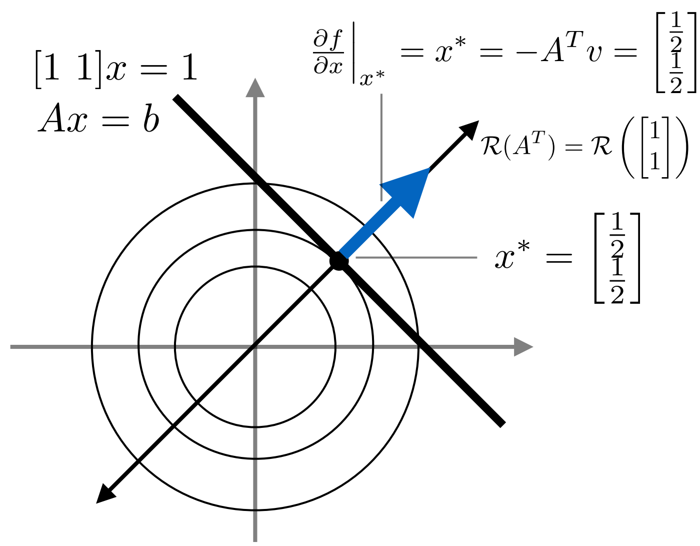
  Solution:

  Plugging in the values we get

$$v^* = -\left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1} = \tfrac{1}{2}, \qquad x^* = -\begin{bmatrix} 1 \\ 1 \end{bmatrix} v^* = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

  Solution:

  The problem is illustrated in the following Figure.



3. **Spherical Level Sets**

   Now consider the optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) = \tfrac{1}{2}|x|_2^2 + c^T x = \tfrac{1}{2}x^T x + c^T x \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

   for $A \in \mathbb{R}^{m \times n}$ full row rank with $m < n$ and $b \in \mathbb{R}^m$. The optimality conditions are given by

$$\frac{\partial f}{\partial x}^T = x + c = -A^T v \tag{3}$$

$$Ax = b \tag{4}$$

   with dual variable $v \in \mathbb{R}^m$. Let x^*, v^* refer to x and v at optimum.

   - **(PTS:0-2)** Solve for v^* in terms of b. (Hint: start by left multiplying (3) by A and substituting in $Ax = b$). Using the solution for v^* solve for x^*.

     Solution:

$$
\begin{aligned}
x + c &= -A^T v \\
Ax + Ac &= -AA^T v \\
b + Ac &= -AA^T v \\
\Rightarrow \qquad v^* &= -(AA^T)^{-1}(b + Ac)
\end{aligned}
$$

$$x^* = -A^T v^* - c = A^T(AA^T)^{-1}(Ac + b) - c$$

- **(PTS:0-2)** Write the objective function in the form from Problem 1.

$$\tfrac{1}{2}x^T x + c^T x = \tfrac{1}{2}z^T z + \text{CONST}$$

for $z = x - \bar{x}$ for some $\bar{x} \in \mathbb{R}^n$. Rewrite the constraint in terms of z, ie. compute $\bar{b}$ such that

$$Ax = b \qquad \Rightarrow \qquad Az = \bar{b}$$

Solution:
Using the form from Problem 1, we have that

$$\tfrac{1}{2}x^T x + c^T x = \tfrac{1}{2}(x + c)^T(x + c) - \tfrac{1}{2}c^T c = \tfrac{1}{2}z^T z - \tfrac{1}{2}c^T c$$

where $z = x - (-c)$. In terms of z, the constraints are given by plugging in $x = z - c$.

$$
\begin{aligned}
Ax = b \qquad \Rightarrow \qquad Az - Ac &= b \\
Az &= b + Ac = \bar{b}
\end{aligned}
$$

- **(PTS:0-2)** Show that

$$z^* = x^* - \bar{x} = A^T(AA^T)^{-1}\bar{b}$$

Solution: Since a constant term in the objective function doesn't affect the optimizer but only the optimal value, we can use the form from Problem 2 to compute the optimum in terms of the variable z.

$$z^* = A^T(AA^T)^{-1}\bar{b}$$

Plugging in the value of $\bar{b}$ gives

$$z^* = A^T(AA^T)^{-1}(Ac + b)$$

as expected.

- **(PTS:0-2)** Consider the above problem for $A = [1\ 1]$ and $b = 1$ and $c^T = [-1\ 1]$ Draw a picture of the optimization space labeling

$$\{x \in \mathbb{R}^2 \mid Ax = b\}, \quad \mathcal{R}(A^T), \quad \text{level sets of } f(x), \quad \bar{x}$$

- **(PTS:0-2)** Also label

$$x^*, \quad z^* = x^* - \bar{x}, \quad -A^T v^*, \quad \frac{\partial f}{\partial x}\Big|_{x^*},$$

and interpret the location of x^* relative to $\bar{x}$

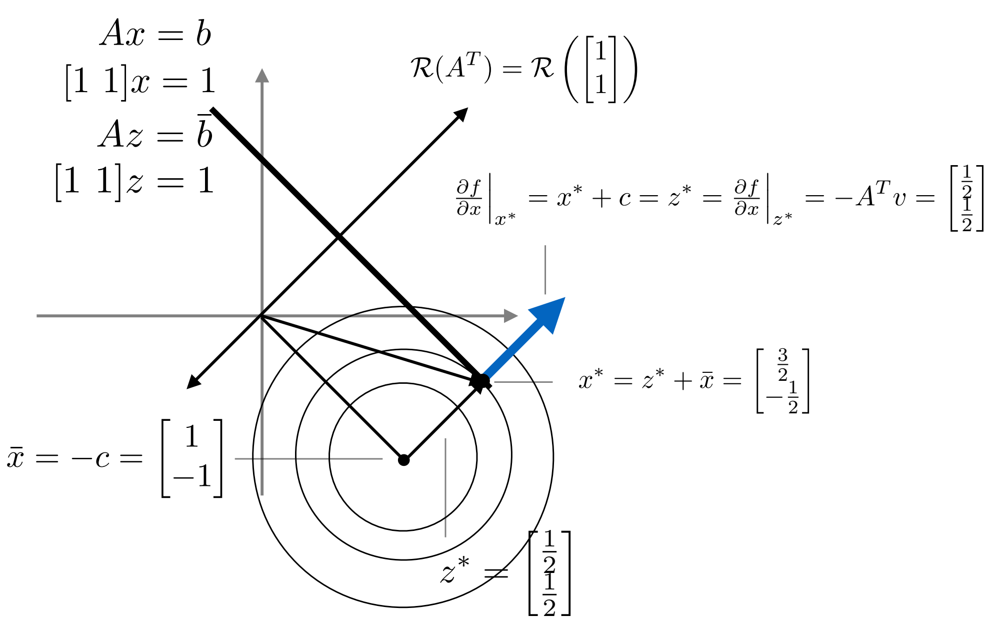
Solution:

The solution is given by

$$x^* = z^* + \bar{x} = z^* - c = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix}$$

The solution is the same as Problem 2 with x^* measured from the center point $\bar{x}$ as opposed to the origin.

The problem is illustrated in the following Figure.



4. **Ellipsoidal Level Sets**

Now consider the optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) = \tfrac{1}{2} x^T Q x + c^T x \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

for $A \in \mathbb{R}^{m \times n}$ full row rank with $m < n$ and $b \in \mathbb{R}^m$. The optimality conditions are given by

$$\frac{\partial f}{\partial x}^T = Qx + c = -A^T v \tag{5}$$

$$Ax = b \tag{6}$$

with dual variable $v \in \mathbb{R}^m$. Let x^*, v^* refer to x and v at optimum.

- **(PTS:0-2)** Solve for v^* in terms of b. (Hint: start by left multiplying (5) by AQ^{-1} and substituting in $Ax = b$). Using the solution for v^* solve for x^*.

  Solution:

$$\begin{aligned} Qx + c &= -A^T v \\ Ax + AQ^{-1}c &= -AQ^{-1}A^T v \\ b + AQ^{-1}c &= -AQ^{-1}A^T v \\ \Rightarrow \qquad v^* &= -(AQ^{-1}A^T)^{-1}(b + AQ^{-1}c) \end{aligned}$$

$$\begin{aligned} x^* &= -Q^{-1}A^T v^* - Q^{-1}c \\ &= Q^{-1}A^T(AQ^{-1}A^T)^{-1}\big(AQ^{-1}c + b\big) - Q^{-1}c \end{aligned}$$

- **(PTS:0-2)** Rewrite the optimization problem using the coordinate transformation $x = Q^{-\frac{1}{2}}x'$ (equivalently $x' = Q^{\frac{1}{2}}x$).

  Solution:

  One way to interpret this problem is that it is the same as optimizing with respect to a function with spherical level sets (Problem 3) in a distorted set of coordinates, namely $x = Q^{-\frac{1}{2}}x'$. Plugging in this new set of coordinates, we get the objective function is

$$f(x) = \tfrac{1}{2}x^T Qx + c^T x = \tfrac{1}{2}(x')^T Q^{-\frac{1}{2}} Q Q^{-\frac{1}{2}} x' + c^T Q^{-\frac{1}{2}} x' = \tfrac{1}{2}(x')^T x' + c^T Q^{-\frac{1}{2}} x' = f(x')$$

  Similarly with the constraints, we can plug in and get

$$Ax = b \qquad \Rightarrow \qquad AQ^{-\frac{1}{2}}x' = b$$

  The optimization problem in the new coordinates is then

$$\begin{aligned} \min_{x'} \quad & \tfrac{1}{2}(x')^T x' + c^T Q^{-\frac{1}{2}} x' \\ \text{s.t.} \quad & AQ^{-\frac{1}{2}}x' = b \end{aligned}$$

- **(PTS:0-2)** Re-solve the optimization problem using the form from Problem 3 in the x' coordinates and show that you get the same solution as your solution above in the x coordinates.

  Solution: Using the form from Problem 3, we get that the solution in the x' coordinates is given by

$$\begin{aligned} x' + Q^{-\frac{1}{2}}c &= -Q^{-\frac{1}{2}}A^T v' \\ AQ^{-\frac{1}{2}}x' + AQ^{-1}c &= -AQ^{-1}A^T v \\ b + AQ^{-1}c &= -AQ^{-1}A^T v' \\ \Rightarrow \qquad (v')^* &= -(AQ^{-1}A^T)^{-1}(b + AQ^{-1}c) \end{aligned}$$

$$x' = -Q^{-\frac{1}{2}}A^T(v')^* - Q^{-\frac{1}{2}}c$$
$$(x')^* = Q^{-\frac{1}{2}}A^T(AQ^{-1}A^T)^{-1}(b + AQ^{-1}c) - Q^{-\frac{1}{2}}c$$
$$\Rightarrow \qquad x^* = Q^{-1}A^T(AQ^{-1}A^T)^{-1}(b + AQ^{-1}c) - Q^{-1}c$$

as expected. Note that $(v')^* = v^*$, ie. the coordinate change on x doesn't affect the value of the Lagrange multipliers or dual variables. The optimal $(x')^*$ gives the same solution as x^* just in the new coordinates.

- **(PTS:0-2)** Consider the above problem for

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad A = [1\ 1], \quad b = 1, \quad c^T = [-1\ 1]$$

Compute the center of the ellipsoidal level sets $\bar{x}$.

Solution:

Using the form from Problem 1, we get that the center of the ellipsoidal level sets is given by

$$\bar{x} = -Q^{-1}c = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -1 \end{bmatrix}$$

- **(PTS:0-2)** Draw a picture of the optimization space labeling

$$\{x \in \mathbb{R}^2 \mid Ax = b\}, \quad \mathcal{R}(A^T), \quad \bar{x}, \quad \text{level sets of } f(x), \quad x^*, \quad -A^T v^*, \quad \frac{\partial f}{\partial x}\Big|_{x^*}$$

Solution: Plugging in the values given gives the solutions...

$$v^* = -1, \qquad x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The optimization problem is then illustrated in the figure below. Note the shape of the level sets of $x^TQx + c^Tx$