EE578B - Convex Optimization - Winter 2021

Homework 4 - Solution

Due Date: Wednesday, Feb 10^{th}, 2021 at 11:59 pm

1. **Constraint Reformulations**

   Consider the optimization problem

$$\begin{aligned} \min_x \quad & f(x) = \tfrac{1}{2}x^T Q x + c^T x \\ \text{s.t.} \quad & x \in \mathcal{X} \end{aligned}$$

   where $\mathcal{X}$ is expressed in two forms.

$$\begin{aligned} \text{Form 1:} \quad & \mathcal{X} = \left\{x \in \mathbb{R}^n \;\middle|\; Ax = b\right\} \\ \text{Form 2:} \quad & \mathcal{X} = \left\{x \in \mathbb{R}^n \;\middle|\; x = Nz + x^0,\ z \in \mathbb{R}^{n-m}\right\} \end{aligned}$$

   for $A \in \mathbb{R}^{m\times n}$ full row rank, $b \in \mathbb{R}^m$, $N \in \mathbb{R}^{n-m}$ where the columns of N form a basis for $\mathcal{N}(A)$ and x^0 satisfies $Ax^0 = b$.

   - **(PTS:0-2)** Solve the optimization problem using the optimality conditions

$$\tfrac{\partial f}{\partial x} = v^T A, \quad Ax = b$$

     for Lagrange multiplier $v \in \mathbb{R}^m$. Solve for x and v.
     Solution:

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

     Plugging in v gives

$$\begin{aligned} Qx + c &= A^T\left(AQ^{-1}A^T\right)^{-1}\Big(b + AQ^{-1}c\Big) \\ x &= Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}\Big(b + AQ^{-1}c\Big) - Q^{-1}c \\ x &= \Big(Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}AQ^{-1} - Q^{-1}\Big)c + Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}b \end{aligned}$$

     For the later parts of the problem it is helpful to note that, $b = AQ^{-1}Qx^0$. We can therefore express x as

$$x = \Big(Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}AQ^{-1} - Q^{-1}\Big)c + Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}AQ^{-1}Qx^0$$

- **(PTS:0-2)** Solve the optimization problem using the optimality conditions

$$\begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial z}\end{bmatrix} = \tau^T[I \ \ -N], \quad x = Nz + x^0$$

for Lagrange multiplier $\tau \in \mathbb{R}^n$. Solve for x, z, and τ.

Solution:

The first optimality condition has two parts.

$$x^TQ + c^T = \tau^T, \qquad 0 = \tau^T N$$

Plugging in $x = Nz + x^0$ and right multiplying by N gives

$$\begin{aligned}
z^TN^TQ + (x^0)^TQ + c^T &= \tau^T \\
z^TN^TQN + (x^0)^TQN + c^TN &= 0 \\
z^T &= -\Big((x^0)^TQN + c^TN\Big)\big(N^TQN\big)^{-1} \\
z &= -\big(N^TQN\big)^{-1}N^T\Big(Qx^0 + c\Big)
\end{aligned}$$

It follows then that

$$\begin{aligned}
x &= Nz + x^0 \\
x &= -N\big(N^TQN\big)^{-1}N^T\Big(Qx^0 + c\Big) + x^0
\end{aligned}$$

- **(PTS:0-2)** Use Form 2 to turn the optimization problem into an unconstrained optimization problem and solve for z.

Solution:

Plugging in the constraint $x = Nz + x^0$ into the objective function gives.

$$\begin{aligned}
f(z) &= \tfrac{1}{2}\Big(Nz + x^0\Big)^TQ\Big(Nz + x^0\Big) + c^T\Big(Nz + x^0\Big) \\
&= \tfrac{1}{2}z^TN^TQNz + \big((x^0)^TQ + c^T\big)Nz + \tfrac{1}{2}(x^0)^TQx^0 + c^Tx^0
\end{aligned}$$

Now optimizing the unconstrained problem gives

$$\begin{aligned}
\tfrac{\partial f}{\partial z} &= z^TN^TQN + \big((x^0)^TQ + c^T\big)N = 0 \\
z &= -(N^TQN)^{-1}N^T\big(Qx^0 + c\big)
\end{aligned}$$

and plugging back for x gives

$$x = -N(N^TQN)^{-1}N^T\big(Qx^0 + c\big) + x^0$$

- **(PTS:0-2)** Show that the answers for all three optimizaton problems are the same. How do x and z relate to each other? How do τ and v relate to each other?

Solution:

The second two solutions are clearly the same. Showing the identity with the first solution requires a little more work. Previously, we constructed the inverse of the matrix

$$\begin{bmatrix} A^T & QN \end{bmatrix}^{-1} = \begin{bmatrix} (AQ^{-1}A^T)^{-1}AQ^{-1} \\ (N^TQN)^{-1}N^T \end{bmatrix}$$

This can be checked by computing

$$\begin{bmatrix} (AQ^{-1}A^T)^{-1}AQ^{-1} \\ (N^TQN)^{-1}N^T \end{bmatrix} \begin{bmatrix} A^T & QN \end{bmatrix} = \begin{bmatrix} (AQ^{-1}A^T)^{-1}AQ^{-1}A^T & (AQ^{-1}A^T)^{-1}AQ^{-1}QN \\ (N^TQN)^{-1}N^TA^T & (N^TQN)^{-1}N^TQN \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

since $AN = 0$. If $BB^{-1} = I$, then $B^{-1}B = I$. Computing this reverse form gives us...

$$\begin{bmatrix} A^T & QN \end{bmatrix} \begin{bmatrix} (AQ^{-1}A^T)^{-1}AQ^{-1} \\ (N^TQN)^{-1}N^T \end{bmatrix} = A^T(AQ^{-1}A^T)^{-1}AQ^{-1} + QN(N^TQN)^{-1}N^T = I$$

Intuitively, this principle is showing that the identity can be written as the sum of two projection matrices onto orthogonal subspaces. This gives us that

$$Q^{-1}A^T(AQ^{-1}A^T)^{-1}AQ^{-1} - Q^{-1} = -N(N^TQN)^{-1}N^T$$

We can plug this in to the first solution to get

$$\begin{aligned} x &= \left(Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}AQ^{-1} - Q^{-1}\right)c + Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}AQ^{-1}Qx^0 \\ x &= -N(N^TQN)^{-1}N^Tc + \left(Q^{-1} - N(N^TQN)^{-1}N^T + \right)Qx^0 \\ x &= -N(N^TQN)^{-1}N^T(c + Qx^0) + x^0 \end{aligned}$$

Relationships:
Clearly $x = Nz + x^0$. The constraint

$$0 = \tau^T N$$

gives that τ is orthgonal to the nullspace of A or equivalently $\tau \in \mathcal{R}(A^T)$, ie. there exists v such that $\tau = A^T v$.

2. **Inequality constraint reformulations**

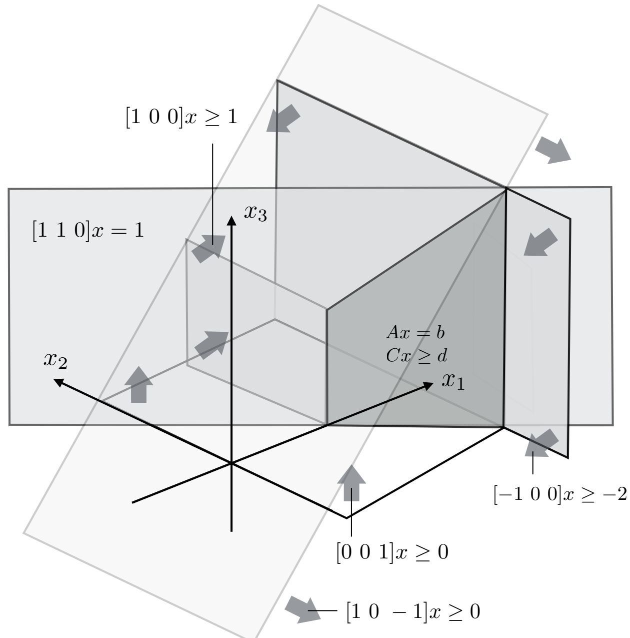
Consider the affine equality and inequality constraints of the form

$$Ax = b, \quad Cx \geq d$$

where $x \in \mathbb{R}^3$ and

$$A = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}, \qquad b = 1$$
$$C = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \qquad d = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

- **(PTS: 0-2)** Plot the set of x's that satisfy these constraints.
  Solution:



- **(PTS: 0-2)** Represent the set $Ax = b$ in the form $x = Nz + x^0$ where the columns of N are a basis for the nullspace of A. Given N and x^0, $Cx \geq d$ implies inequality constraints on z. Write down these constraints, ie. find E and h such that $Ez \geq h$. If you chose a different N and/or x^0 would these constraints be the same?

  Solution:

  The set $Ax = b$ can be transformed to $x = Nz + x^0$ where

  $$N = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad x^0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

  Plugging this into $Cx \geq d$ gives

  $$\begin{aligned} Cx &\geq d \\ Ez = CNz &\geq d - Cx^0 = h \end{aligned}$$

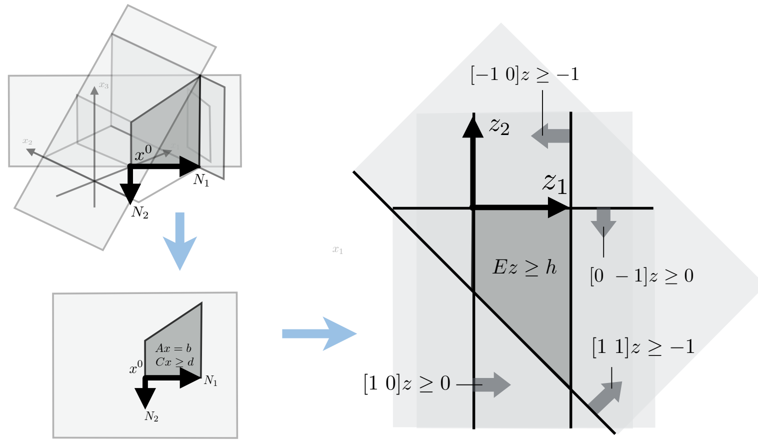
  For the N and x^0, we chose we get

  $$E = CN = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix},$$

  $$h = d - Cx^0 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

  A different N and x^0 would provide a different parametrization of the z space and thus the constraints on z would be different.

- **(PTS: 0-2)** Plot the set of z's such that $Ez \geq h$.



3. **Inequality Constraints and Positive Matrices (EXTRA CREDIT) (PTS: 0-2)**

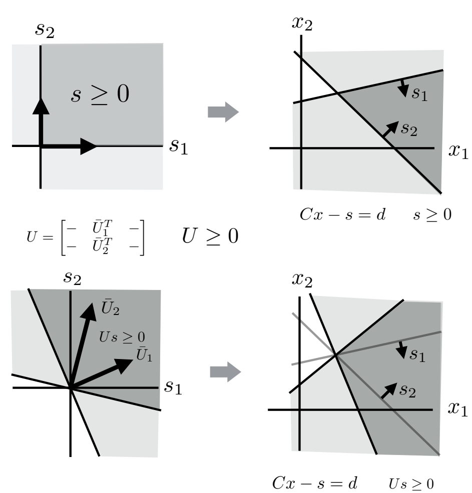
   The inequality constraint $Cx \geq d$ for $C \in \mathbb{R}^{p \times n}$ and $d \in \mathbb{R}^p$ can be rewritten as

$$Cx \geq d \qquad \iff \qquad Cx = d + s, \quad s \geq 0$$

   using a slack variable $s \in \mathbb{R}^p$. How does the constrained set change if $s \geq 0$ is replaced with the constraint $Us \geq 0$ for some matrix U with positive elements. Suggestion: experiment with an example where $C \in \mathbb{R}^{2 \times 2}$.

   Solution:

   For a positive matrix U, the set $Us \geq 0$ gets larger and includes the positive orthant, ie. $s \geq 0 \Rightarrow Us \geq 0$. Algebraically, this is because the sum of positive elements is positive. Geometrically, when the rows of U point in the interior of the positive orthant, some points with some negative values of s still satisfy $Us \geq 0$. This same expansion happens to the set of x's that satisfy $Cx - s = d$ as shown in the Figure



4. **Complementary Slackness**

   For $x \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^n$

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

$$\mu_i x_i = 0,\ x_i \geq 0,\ \mu_i \geq 0, \qquad \Longleftrightarrow \qquad \mu_i = 0 \ \text{ OR } \ x_i = 0$$

  Solution: This problem is to experiment with the on/off nature of the complementary slackness constraint.

  If $\mu_i = 0$ or $x_i = 0$ then $\mu_i x_i = 0$. Conversely, if $\mu_i x_i = 0$ and $\mu_i > 0$ then $x_i = 0$ and if $\mu_i x_i = 0$ and $x_i > 0$, then $\mu_i = 0$.

- **(PTS: 0-2)** Show that if $x_i \geq 0$ and $\mu_i \geq 0$ for all i then

$$\mu_i x_i = 0,\ \forall i \qquad \Longleftrightarrow \qquad \mu^T x = 0$$

  Solution:

  If $\mu_i = 0$ and $x_i = 0$ for all i, then summing gives $\mu^T x = 0$. If $\mu^T x = 0$ and there is any i such that $\mu_i x_i > 0$, then $\sum_{j \neq i} \mu_j x_j < 0$ which is impossible since each term in the sum is positive.

5. **Lagrangians**

Consider the optimization problem

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

where

$$A = \begin{bmatrix} - & \bar{A}_1^T & - \\ - & \bar{A}_2^T & - \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.9 \end{bmatrix} \tag{1}$$

- **(PTS:0-2)** Write the Lagrangian for the optimization problem $\mathcal{L}(x, v)$ with a dual variable $v \in \mathbb{R}^2$. (You can leave your answer in terms of A and b.)

  Solution:

$$\begin{aligned} \mathcal{L}(x, v) &= f(x) + v^T (Ax - b) \\ &= \tfrac{1}{2} x^T x + v^T (Ax - b) \end{aligned}$$

  Note that the sign of v doesn't matter since it is an equality constraint.

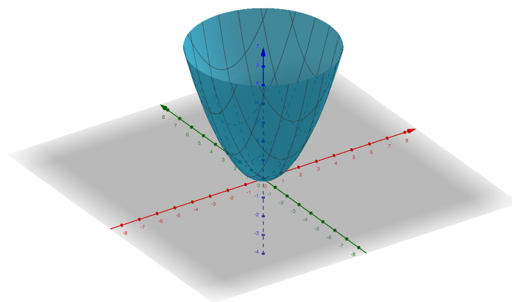
- **(PTS:0-2)** Plot each of the following surfaces $h(x)$ in the $[x_1 \ x_2 \ h(x)]^T$ space. For the expressions that include v_1 and v_2, hold v_1 and v_2 fixed for each plot.

$$h(x) = f(x), \quad h(x) = v_1(\bar{A}_1^T x - b_1), \quad h(x) = v_2(\bar{A}_2^T x - b_2), \quad h(x) = \mathcal{L}(x, v)$$

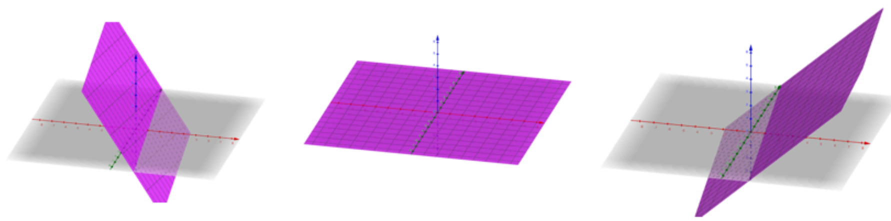
  Now vary the values of v_1 and v_2. How does each surface change for positive values of v_1 and v_2? What about for negative values?

  Solution:

  Plotting, we see $f(x)$

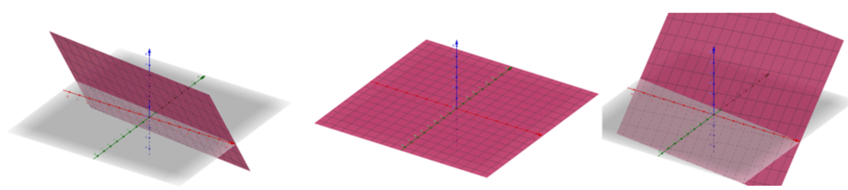


$v_1(\bar{A}_1^T x - b_1)$ for different values of v_1



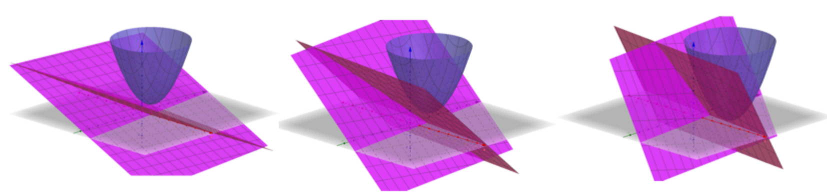
$v_1 < 0$ $\quad$ $v_1 = 0$ $\quad$ $v_1 > 0$

$v_2(\bar{A}_2^T x - b_2)$ for different values of v_2,

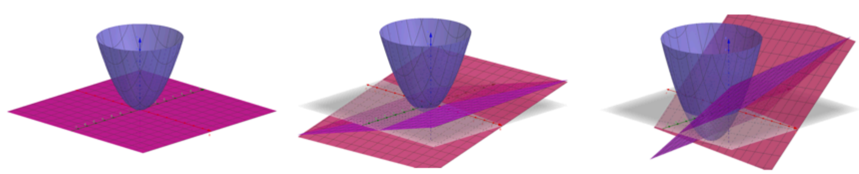


$v_2 < 0$ $\quad$ $v_2 = 0$ $\quad$ $v_2 > 0$

and L



$v_1 = v_2 = -0.5$ $\quad$ $v_1 = v_2 = -1.$ $\quad$ $v_1 = v_2 = -2$



$v_1 = v_2 = 0.0$ $\quad$ $v_1 = v_2 = 0.3$ $\quad$ $v_1 = v_2 = 1$

Note that the constraint surfaces flip over as v_i crosses 0, and that there is some v that maximizes the minimum value of $\mathcal{L}$ (the bottom of the bowl).

- **(PTS:0-2)** Now replace $Ax = b$ with $Ax \geq b$. Does the Lagrangian change? What new constraints do we have to impose on the dual variables v_1 and v_2? How does this affect how the surfaces

$$h(x) = v_1(\bar{A}_1^T x - b_1), \quad h(x) = v_2(\bar{A}_2^T x - b_2), \quad h(x) = \mathcal{L}(x, v)$$

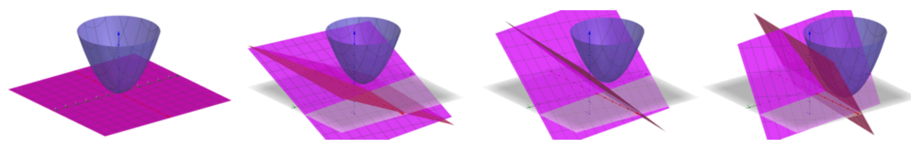
change with v_1 and v_2?

Solution:

The sign of the dual variable in the Lagrangian now matters and v must be $v \geq 0$. The Lagrangian is given by

$$\mathcal{L}(x, v) = \frac{1}{2}x^T x - v^T(Ax - b)$$

The dual variables are trying to maximize the Lagrangian so the negative sign ensures that it is in the primal variables interests to make $Ax - b \geq 0$ so that any positive value of v can only make the Lagrangian smaller. The new positivity constraints on the dual variables make it so that the surfaces can't dip below zero on one side. The Lagrangian surface for different values of $v \geq 0$ is shown below



$v_1 = v_2 = 0$ $\qquad$ $v_1 = v_2 = 0.5$ $\qquad$ $v_1 = v_2 = 1$ $\qquad$ $v_1 = v_2 = 2$

Note that in this case, v still maximizes the minimum of the Lagrangian by forcing x to be at the intersection of the two linear boundaries.