NTNU

Kunnskap for en bedre verden

TTK4135 - Optimization and Control

Exercise #4

Author:
Sondre Pedersen

February 13, 2025

Problem 1: Quadratic programming

(a) Under which conditions is a QP-problem convex? Why is convexity an important property?

(b) Go through the proof of Theorem 16.2. If the lemma $Z^\top GZ \geq 0$ instead of $Z^\top GZ > 0$, how would the wording of the proof change?

(c) Based on Example 16.4, show how Algorithm 16.3 finds the solution if the starting point is $x = \begin{bmatrix}2 & 0\end{bmatrix}^\top$ and we assume that only the constraint $-x_1 + 2x_2 + 2 \geq 0$ is active. This means $\mathcal{W}_0 = \{3\}$.

(d) Define the dual problem for the QP-problem in Example 16.4.

(e) Explain how the dual optimization problem can be used to give an over-estimate of $q(\overline{x}) - q(x^*)$, when x^* is not known.

(a) Convexity

Convexity is an important property because local solutions are guaranteed to be global solutions in convex optimization problems. They are simply much easier to solve.
QP-problems are convex when the Hessian matrix G is positive semidefinite.

(b) Proof

If we assume $Z^\top GZ > 0$, this means the quadratic form $(Z^\top GZ)$ is positive definite, ensuring that $q(x) > q(x^*)$ for any $x \neq x^*$, which guarantees that x^* is a strict local minimizer.
On the other hand, if we change the assumption to $Z^\top GZ \geq 0$, we are now only assuming positive semidefiniteness. This means that some directions in the same space might have zero curvature, allowing $q(x)$ to be the same as $q(x^*)$ for certain small changes. This is a non-strict local minimum, $q(x) \geq q(x^*)$ but there

(c) Solution

The problem on matrix form is

$$\min_x \qquad \frac{1}{2}x^\top Gx + c^\top x \qquad \text{s.t.} \qquad Ax = b.$$

where

$$G = \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix},\ c = \begin{bmatrix}-2\\-5\end{bmatrix},\ A = \begin{bmatrix}1 & -2\\-1 & -2\\-1 & 2\\1 & 0\\0 & 1\end{bmatrix},\ b = \begin{bmatrix}-2\\-6\\-2\\0\\0\end{bmatrix}.$$

Iteration k = 0

- $x^{(0)} = \begin{bmatrix}2 & 0\end{bmatrix}^\top$
- $\mathcal{W}^{(0)} = \{3\}$
- $A^{(0)} = \begin{bmatrix}-1 & 2\end{bmatrix}$
- $b^{(0)} = 0$
- $g^{(0)\top} = Gx^{(0)} + c = \begin{bmatrix}2 & -5\end{bmatrix}^\top$

Calculate x and lambda (dropping superscript on A and b here):

$$p^{(0)} = -G^{-1}(I - A^\top(AG^{-1}A^\top)^{-1}AG^{-1})c + G^{-1}A^\top(AG^{-1}A^\top)^{-1}b = \begin{bmatrix}0.2\\0.1\end{bmatrix}$$
$$\lambda^{(0)} = -(AG^{-1}A^\top)^{-1}(AG^{-1}c + Ax) = -2.4$$

Now,

$$x^{(1)} = x^{(0)} + p^{(0)} = \begin{bmatrix}2\\0\end{bmatrix} + \begin{bmatrix}0.2\\0.1\end{bmatrix} = \begin{bmatrix}2.2\\0.1\end{bmatrix}.$$

which is a feasible point.

Iteration k = 1

- $g^{(1)} = Gx^{(1)} + c = \begin{bmatrix} 2.4 & -4.8 \end{bmatrix}^\top$

$$
\begin{aligned}
p^{(1)} &= -G^{-1}(I - A^\top (AG^{-1}A^\top)^{-1} AG^{-1})c + G^{-1}A^\top (AG^{-1}A^\top)^{-1} b \approx \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\lambda^{(1)} &= -(AG^{-1}A^\top)^{-1}(AG^{-1}c + Ax) = -2.4
\end{aligned}
$$

Iteration k = 2

- $x^{(2)} = x^{(1)} = \begin{bmatrix} 2.2 & 0.1 \end{bmatrix}^\top$
- $\mathcal{W}^{(2)} = \emptyset$
- $g^{(2)\top} = Gx^{(2)} + c = \begin{bmatrix} 2.4 & -4.8 \end{bmatrix}^\top$

$$
p^{(2)} = -G^{-1}(I - A^\top (AG^{-1}A^\top)^{-1} AG^{-1})c + G^{-1}A^\top (AG^{-1}A^\top)^{-1} b = \begin{bmatrix} -1.2 \\ 2.4 \end{bmatrix}
$$

This is not a valid point. The blocking constraint is constraint 1, with $\alpha_2 = \frac{2}{3}$.

Iteration k = 3

- $x^{(3)} = x^{(1)} + \alpha_2 p_2 = \begin{bmatrix} 1.4 & 1.7 \end{bmatrix}^\top$
- $\mathcal{W}^{(3)} = \{1\}$
- $A^{(3)} = \begin{bmatrix} 1 & -2 \end{bmatrix}$
- $b^{(3)} = -2$
- $g^{(3)\top} = Gx^{(3)} + c = \begin{bmatrix} 0.8 & -1.6 \end{bmatrix}^\top$

$$
\begin{aligned}
p^{(3)} &= -G^{-1}(I - A^\top (AG^{-1}A^\top)^{-1} AG^{-1})c + G^{-1}A^\top (AG^{-1}A^\top)^{-1} b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\lambda^{(3)} &= -(AG^{-1}A^\top)^{-1}(AG^{-1}c + Ax) = 0.8
\end{aligned}
$$

So we can conclude that this is the answer.

(d) The dual

The values are

$$
\begin{aligned}
\min_x \quad & q(x) = \frac{1}{2} x^\top G x + x^\top c \\
\text{s.t.} \quad & Ax - b \geq 0
\end{aligned}
$$

where

$$
G = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \; A = \begin{bmatrix} 1 & -2 \\ -1 & -2 \\ -1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \; c = \begin{bmatrix} -2 \\ -5 \end{bmatrix}, \; b = \begin{bmatrix} -2 \\ -6 \\ -2 \\ 0 \\ 0 \end{bmatrix}.
$$

The dual becomes

$$
\begin{aligned}
\max_{x,\lambda} \quad & \frac{1}{2} x^\top G x + x^\top c - \lambda^\top (Ax - b) \\
\text{s.t.} \quad & Gx + c - A^\top \lambda = 0 \\
& \lambda \geq 0
\end{aligned}
$$

(e) Dual as over-estimate

The dual optimization problem can be used to give an over-estimate of $q(\overline{x}) - q(x^*)$ when x^* is not known. Any feasible $\overline{x}$ and any $\overline{\lambda} \geq 0$ will give $f(\overline{\lambda}) \leq q(x^*)$. Therefore,

$$q(\overline{x}) - q(x^*) \leq q(\overline{x}) - f(\overline{\lambda}).$$

Problem 1: Production Planning and Quadratic Programming

Two reactors R_1 and R_2 produce two products A and B. To make 1000kg of A, 2 hours of R_1 and 1 hour of R_2 are required. To make 1000kg of B, 1 hour of R_1 and 3 hours of R_2 are required. The order of R_1 and R_2 does not matter. R_1 and R_2 are available for 8 and 15 hours, respectively. We want to maximize the profit from selling the two products.
The profit now depends on the production rate:

- the profit from A is $3 - 0.4x_1$ per tonne produced,
- the profit from B is $2 - 0.2x_2$ per tonne produced,

where x_1 is the production of product A and x_2 is the production of product B (both in number of tonnes).

(a) Formulate this as a quadratic program.

(b) Make a contour plot and sktech constraints.

(c) Find the production of A and B that maximizes the total profit.

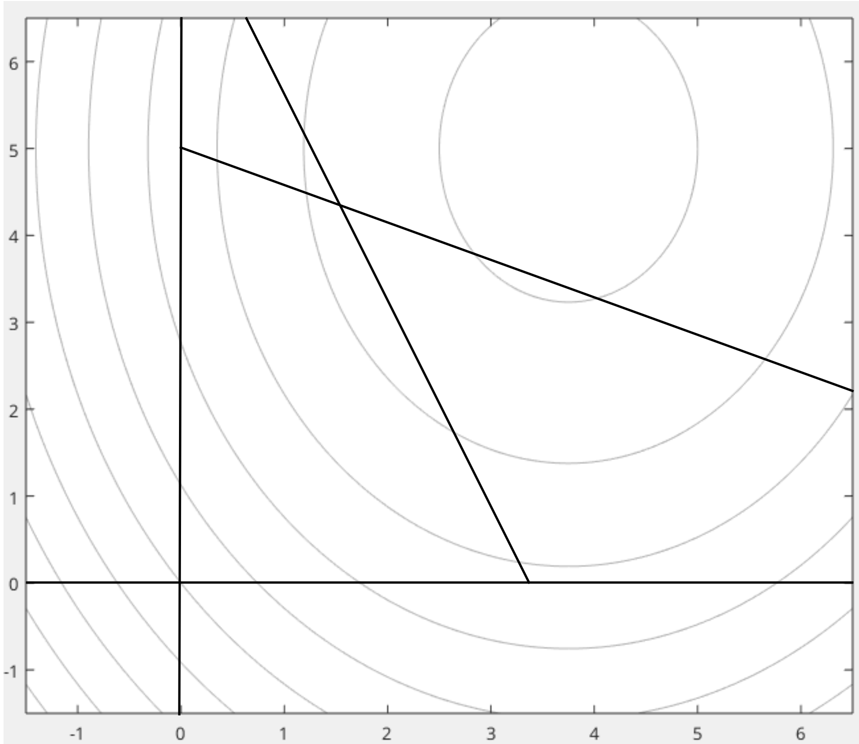
(d) The solution is calculated by an active-set method. Explain how the method works based on the sequence of iterations from (c).

(e) Compare the solution in (c) with Problem 2 (c) in Exercise 3 and comment.

$$f(x) = -[(3-0.4x_1)x_1 + (2-0.2x_2)x_2] = -(3-0.4x_1)x_1 - (2-0.2x_2)x_2 = \frac{1}{2}x^\top \begin{bmatrix} 0.8 & 0 \\ 0 & 0.4 \end{bmatrix} x + \begin{bmatrix} -3 & -2 \end{bmatrix} x.$$

$$\begin{aligned}
\min_{x} \quad & q(x) = \frac{1}{2}x^\top G x + x^\top c \\
\text{s.t.} \quad & -2x_1 - x_2 \geq -8 \\
& -x_1 - 3x_2 \geq -15 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}$$

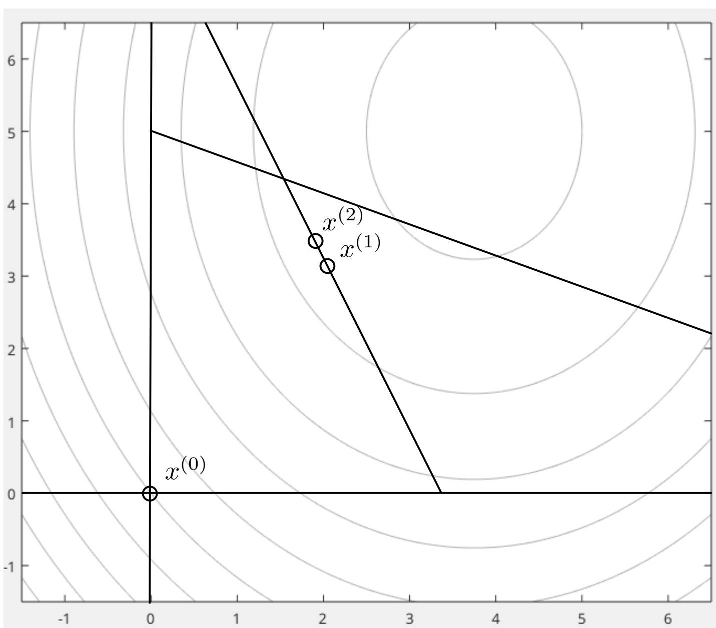
(b) Contour



(c) Solution

Plugging the values from (a) into the script gave these outputs:

- $\begin{bmatrix} x_1^{(0)} & x_2^{(0)} \end{bmatrix}^\top = \begin{bmatrix} 0 & 0 \end{bmatrix}^\top$
- $\begin{bmatrix} x_1^{(1)} & x_2^{(1)} \end{bmatrix}^\top = \begin{bmatrix} 2.4 & 3.2 \end{bmatrix}^\top$
- $\begin{bmatrix} x_1^{(2)} & x_2^{(2)} \end{bmatrix}^\top = \begin{bmatrix} 2.25 & 3.5 \end{bmatrix}^\top$



The solution is not found at an intersection. Only one constraint is active.

(d) Active-set method

Is an iterative approach to solving QP problems by choosing a subset of the constraints to be active. This gives us a reduced, equality constrained problem that we can solve. This solution can be checked for optimality, and if it is not optimal, add or remove constraints from the active set before repeating the process.

(e) Compare with LP

The solution in exercise 3 used the fact that an optimal solution had to lie in an intersection, and could therefore iterate over those. Here we cannot make that assumption and must therefore use a completely different method for solving the problem. We see that the second and third points differ.