

Name:

Present group members:

Worksheet 10-2: Q1 Consider the following problem:

$$\begin{aligned} \min_{x,y} \quad & (x - 0.1)^2 + y^2 \\ \text{s.t.} \quad & x + y \leq 1, \\ & x - 2y \leq 0. \end{aligned}$$

We will find the KKT conditions

- (a) Write the Lagrangian function of this problem.

$$L(x, y, \lambda_1, \lambda_2) = (x - 0.1)^2 + y^2 + \lambda_1(x + y - 1) + \lambda_2(x - 2y).$$

- (b) Write down the constraints from the KKT condition (stationarity condition) for the problem.

$$\nabla_{\mathbf{x}}L(x, y, \lambda_1, \lambda_2) = \begin{bmatrix} 2(x - 0.1) + \lambda_1 + \lambda_2 \\ 2y + \lambda_1 - 2\lambda_2 \end{bmatrix} = \mathbf{0}.$$

So the resulting equations are

$$\begin{aligned} 2(x - 0.1) + \lambda_1 + \lambda_2 &= 0 \\ 2y + \lambda_1 - 2\lambda_2 &= 0 \end{aligned}$$

- (c) Write down the complementary slackness conditions for this problem.

$$\begin{aligned} \lambda_1(x + y - 1) &= 0 \\ \lambda_2(x - 2y) &= 0 \end{aligned}$$

- (d) Write down the feasibility constraints (including the non-negativity constraints on the Lagrange multipliers) for this problem.

$$\begin{aligned} x + y - 1 &\leq 0 \\ x - 2y &\leq 0 \\ \lambda_1 &\geq 0 \\ \lambda_2 &\geq 0 \end{aligned}$$

- (e) Count the equations/inequalities found in the previous three parts. You should have 8 of them.
- (f) Do you expect the local optima (if any) to be global optima? Why?

Yes, the function is convex, and the constraints are linear (so also convex), thus the problem is convex. The solution to the KKT requirements will be a global optimum.

(g) In theory, you could now solve for the unknowns $x, y, \lambda_1, \lambda_2$ using the equations/inequalities you have. This turns into checking cases:

- Case 1: $\lambda_1 = \lambda_2 = 0$.
- Case 2: $\lambda_1 = 0, \lambda_2 > 0$.
- Case 3: $\lambda_1 > 0, \lambda_2 = 0$.
- Case 4: $\lambda_1 > 0, \lambda_2 > 0$.

In this example, Cases 3, and 4 do not lead to feasible points. Find the KKT point(s) for Cases 1 and 2 if they exist.

- *Case 1*

- We are assuming $\lambda_1 = \lambda_2 = 0$, so the stationarity conditions become

$$\begin{aligned} 2(x - 0.1) &= 0 \\ 2y &= 0 \end{aligned}$$

- Solving these equations gives us $x = 0.1, y = 0$.
- Now we need to check if this solution is feasible. We do this by checking the feasibility constraints:

$$\begin{aligned} x + y - 1 &= 0.1 + 0 - 1 = -0.9 \leq 0 \quad \checkmark \\ x - 2y &= 0.1 - 2 \cdot 0 = 0.1 \leq 0 \quad \times \\ \lambda_1 &= 0 \geq 0 \quad \checkmark \\ \lambda_2 &= 0 \geq 0 \quad \checkmark \end{aligned}$$

- Since the second constraint is not satisfied, this solution is not feasible, so it cannot be a KKT point.

- *Case 2*

- We are assuming $\lambda_1 = 0, \lambda_2 > 0$. So the stationarity conditions become

$$\begin{aligned} 2(x - 0.1) + \lambda_2 &= 0 \\ 2y - 2\lambda_2 &= 0 \end{aligned}$$

- We can eliminate λ_2 from these equations to obtain

$$2(x - 0.1) + y = 0$$

- Since $\lambda_2 > 0$ (specifically $\lambda_2 \neq 0$), the complementary slackness condition $\lambda_2(x - 2y) = 0$ implies that $x - 2y = 0$. So $y = \frac{x}{2}$.
- Solving this equation together with $2(x - 0.1) + y = 0$ gives us $x = 0.08, y = 0.04$.
- Using $2y - 2\lambda_2 = 0$, we have that $\lambda_2 = 0.04$.
- Now we need to check if this solution is feasible. We do this by checking the feasibility constraints:

$$\begin{aligned} x + y - 1 &= 0.08 + 0.04 - 1 = -0.88 \leq 0 \quad \checkmark \\ x - 2y &= 0.08 - 2 \cdot 0.04 = 0 \leq 0 \quad \checkmark \\ \lambda_1 &= 0 \geq 0 \quad \checkmark \\ \lambda_2 &= 0.04 > 0 \geq 0 \quad \checkmark \end{aligned}$$

(h) What can you conclude about the optimal solution of this problem?

We have a unique KKT point, and the problem is convex. This means that the minimum is achieved at $(0.08, 0.04)$.

(i) Which of the constraints are active at the optimal solution?

- *The second constraint, $x - 2y \leq 0$ is active, since $x - 2y = 0.08 - 2 \cdot 0.04 = 0$ means it holds with equality at the optimal solution.*
- *The first constraint, $x + y \leq 1$, is not active, since $x + y = 0.08 + 0.04 = 0.12 < 1$, so it does not hold with equality at the optimal solution.*

(j) Take a look at [this desmos plot](#). Visually confirm that the solution you found above is indeed the optimal solution. Then turn on the plot of $g(x, y)$, which is a copy of $f(x, y)$. Change the constraints for g (the portions inside $\{\dots\}$). Which constraint can be changed without changing the optimal solution? Which constraint can be changed to change the optimal solution? What does this have to do with the previous question?

- *You can change the first constraint $x + y \leq 1$ without changing the optimal solution, as long as the constraint is not active at the optimal solution. For example, if we change the first constraint to $x + y \leq 0.5$, the optimal solution remains unchanged.*
- *If we change the second constraint $x - 2y \leq 0$ to $x - 2y \leq -0.01$, then the optimal solution changes, since the new constraint is active at the optimal solution.*
- *This can be seen in [this version of the desmos plot](#). The function $g(x, y)$ has the $x - 2y \leq 0$ constraint changed and the optimal solution changes. The function $h(x, y)$ has the $x + y \leq 1$ constraint changed and the optimal solution does not change.*

Worksheet 10-2: Q2

Consider the optimization problem $\min\{\mathbf{x}^T Q \mathbf{x} + 2\mathbf{c}^T \mathbf{x} : A\mathbf{x} = \mathbf{b}\}$, where $Q \in \mathbb{R}^{n \times n}$ is a positive definite matrix, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, and A is an $m \times n$ matrix with linearly dependent rows. This is a convex optimization problem. Answer the following:

- (a) Find the Lagrangian as a single matrix multiplication equation.

The Lagrangian general equation is

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}).$$

For a linearly constrained system, this equation simplifies to

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T (A\mathbf{x} - \mathbf{b}) + \boldsymbol{\mu}^T (C\mathbf{x} - \mathbf{d}).$$

For this problem we only have a set of m equality constraints, so the Lagrangian for this problem reads

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\mu}) &= f(\mathbf{x}) + \boldsymbol{\mu}^T (A\mathbf{x} - \mathbf{b}) \\ &= \mathbf{x}^T Q \mathbf{x} + 2\mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (A\mathbf{x} - \mathbf{b}) \\ &= \mathbf{x}^T Q \mathbf{x} + 2\mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T A \mathbf{x} - \boldsymbol{\mu}^T \mathbf{b} \end{aligned}$$

Note how we adapted the general form to our problem by noting that the matrix C and \mathbf{d} in the general form are replaced by the matrix A and vector \mathbf{b} from our problem. We have no inequality constraints, so the terms containing $\boldsymbol{\lambda}$ are omitted.

- (b) Noting that we treat vectors as column vectors, write down the dimensions of all the matrices and vectors in the Lagrangian above.

- \mathbf{x} is $n \times 1$.
- Q is $n \times n$.
- \mathbf{c} is $n \times 1$.
- A is $m \times n$.
- \mathbf{b} is $m \times 1$.
- $\boldsymbol{\mu}$ is $m \times 1$.

(c) For a column vector \mathbf{b} , $\nabla_{\mathbf{x}}(\mathbf{b}^T \mathbf{x}) = \mathbf{b}$. Use this to determine $\nabla_{\mathbf{x}}(\boldsymbol{\mu}^T A \mathbf{x})$?

- $\boldsymbol{\mu}^T A$ is a row vector, ($\boldsymbol{\mu}$ is $m \times 1$ and A is $m \times n$ means $\boldsymbol{\mu}^T A$ is $1 \times n$).
- So we can write $\boldsymbol{\mu}^T A \mathbf{x}$ as $\mathbf{b}^T \mathbf{x}$ where $\mathbf{b}^T = \boldsymbol{\mu}^T A$.
- Therefore,

$$\nabla_{\mathbf{x}}(\boldsymbol{\mu}^T A \mathbf{x}) = (\boldsymbol{\mu}^T A)^T = A^T \boldsymbol{\mu}.$$

(d) Using the same as above, what is $\nabla_{\mathbf{x}}(2\mathbf{c}^T \mathbf{x})$?

- $2\mathbf{c}^T \mathbf{x}$ is a scalar, since \mathbf{c}^T is $1 \times n$ and \mathbf{x} is $n \times 1$.
- We can write $2\mathbf{c}^T \mathbf{x}$ as $\mathbf{b}^T \mathbf{x}$ where $\mathbf{b}^T = 2\mathbf{c}^T$.
- Therefore, $\nabla_{\mathbf{x}}(2\mathbf{c}^T \mathbf{x}) = (2\mathbf{c}^T)^T = 2\mathbf{c}$.

(e) Since Q is symmetric, $\nabla_{\mathbf{x}}(\mathbf{x}^T Q \mathbf{x}) = 2Q\mathbf{x}$. What is $\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\mu})$?

Using the parts above, we have

$$\begin{aligned}\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\mu}) &= \nabla_{\mathbf{x}}(\mathbf{x}^T Q \mathbf{x}) + \nabla_{\mathbf{x}}(2\mathbf{c}^T \mathbf{x}) + \nabla_{\mathbf{x}}(\boldsymbol{\mu}^T A \mathbf{x}) - \nabla_{\mathbf{x}}(\boldsymbol{\mu}^T \mathbf{b}) \\ &= 2Q\mathbf{x} + 2\mathbf{c} + A^T \boldsymbol{\mu} - \mathbf{0} \\ &= 2Q\mathbf{x} + 2\mathbf{c} + A^T \boldsymbol{\mu}\end{aligned}$$

(f) Write down the KKT conditions (stationarity and feasibility), and to simplify down the road, replace $\boldsymbol{\mu} = 2\boldsymbol{\gamma}$. What are the unknowns in these equations?

The KKT conditions are

- *The original **Stationarity**:*

$$\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\mu}) = 2Q\mathbf{x} + 2\mathbf{c} + A^T \boldsymbol{\mu} = \mathbf{0}$$

- *The modified **Stationarity** after scaling the Lagrange multipliers:*

$$\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\mu}) = 2(Q\mathbf{x} + \mathbf{c} + A^T \boldsymbol{\gamma}) = \mathbf{0}$$

where $\boldsymbol{\mu} = 2\boldsymbol{\gamma}$. This can be simplified to

$$Q\mathbf{x} + \mathbf{c} + A^T \boldsymbol{\gamma} = \mathbf{0}$$

- ***Feasibility**: $A\mathbf{x} = \mathbf{b}$.*

The unknowns in this equation are \mathbf{x} which contains n elements, and the vector $\boldsymbol{\gamma}$, which contains m elements (same number as the rows of the matrix A). So the total number of unknowns is $m + n$.

(g) Solve the stationarity condition for \mathbf{x} .

Solving the stationarity condition for \mathbf{x} gives

$$\mathbf{x} = -Q^{-1}(\mathbf{c} + A^T \boldsymbol{\gamma}).$$

(h) Substitute the expression for \mathbf{x} into the feasibility constraint, and solve for any other unknown.

Substituting the expression for \mathbf{x} into the feasibility constraint we obtain

$$-AQ^{-1}(\mathbf{c} + A^T \boldsymbol{\gamma}) = \mathbf{b},$$

Which we can solve for $\boldsymbol{\gamma}$

$$\boldsymbol{\gamma} = -(AQ^{-1}A^T)^{-1}(\mathbf{b} + AQ^{-1}\mathbf{c}).$$

(i) What is the stationary point in terms of \mathbf{c} , Q , A , and \mathbf{b} ?

The optimal solution is the one we found in (g), with the $\boldsymbol{\gamma}$ value found in (h),

$$\mathbf{x} = -Q^{-1}(\mathbf{c} - A^T(AQ^{-1}A^T)^{-1}(\mathbf{b} + AQ^{-1}\mathbf{c})).$$

(j) Is the stationary point optimal? Explain.

Yes. The matrix Q in the quadratic form is positive definite, so the objective function is strictly convex. The constraint is affine, so convex. Therefore, this is a convex optimization problem, and KKT points are global optima. In this case, we have a unique KKT point so it is the unique global optimum.