A4Q1: KKT Conditions and Convex Problems

20273229

November 14, 2023

1. (a) Substituting the property function into the Lagrange formula, we have

$$\mathcal{L}(\vec{w}, \lambda) = \frac{1}{2}[\vec{w} - \vec{g}]^T[\vec{w} - \vec{g}] + \lambda(([1\ 0]\vec{w})^2 - 4)$$

The derivative of the Lagrange function with respect to w_1 is

$$\frac{\partial \mathcal{L}}{\partial w_1} = w_1 - g_1 + 2\lambda w_1$$

The derivative of the Lagrange function with respect to w_2 is

$$\frac{\partial \mathcal{L}}{\partial w_1} = w_2 - g_2$$

Since $g_1 = g_2 = 3$, we know $w_2 - 3 = 0$. Thus, $w_2 = 3$.
The derivative of the Lagrange function with respect to λ is

$$\frac{\partial \mathcal{L}}{\partial \lambda} = ([1\ 0]\vec{w})^2 - 4$$

By setting these to 0, we can solve for $\hat{w}$ and λ.

$$\begin{aligned} 0 &= ([1\ 0]\hat{w})^2 - 4 \\ \sqrt{4} &= ([1\ 0]\hat{w}) \\ \hat{w}_1 &= \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \hat{w}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix} \end{aligned}$$

Substitute these into the partial derivative of the Lagrange function with respect to w_1 to find λ and with $g = [3; 3]$. First, using $\hat{w}_1$

$$\begin{aligned} 0 &= 2 - 3 + 2\lambda_1(2) \\ \lambda_1 &= 0.25 \end{aligned}$$

Then, using $\hat{w}_2$

$$\begin{aligned} 0 &= -2 - 3 + 2\lambda_2(-2) \\ \lambda_2 &= -1.25 \end{aligned}$$

Based on the dual feasibility condition, λ_2 is negative; thus, $\hat{w}_2$ is not a valid solution. Verify that λ_1 and $\hat{w}_1$ are valid solutions using the KKT conditions. Primal feasibility:

$$\begin{aligned} A\hat{w}_1 &\leq 4 \\ 2 &\leq 4 \end{aligned}$$

Dual feasibility:

$$\lambda_1 \geq 0$$
$$0.25 \geq 0$$

Stationary:

$$\begin{bmatrix} \hat{w}_1 - g_1 + 2\lambda\hat{w}_1 \\ \hat{w}_2 - g_2 \end{bmatrix} = \vec{0}$$
$$\begin{bmatrix} 2 - 3 + 2(0.25)(2) \\ 3 - 3 \end{bmatrix} = \vec{0}$$

Complementary slackness:

$$\lambda_1(A\hat{w}_1 - b) = 0$$
$$0.25(2^2 - 4) = 0$$

All conditions have been met. Therefore, $\hat{w} = [2; 3]$ and $\hat{\lambda} = 0.25$. Since $\hat{\lambda}$ is positive, we can say that the constraint is active and the minimizer is on the boundary of the inequality constraint.

(b) Substituting the property function into the Lagrange formula, we have

$$\mathcal{L}(\vec{w}, \lambda) = \frac{1}{2}[\vec{w} - \vec{g}]^T[\vec{w} - \vec{g}] + \lambda(([0\ 1]\vec{w})^2 - 4)$$

The derivative of the Lagrange function with respect to w_1 is

$$\frac{\partial \mathcal{L}}{\partial w_1} = w_1 - g_1$$

The derivative of the Lagrange function with respect to w_2 is

$$\frac{\partial \mathcal{L}}{\partial w_2} = w_2 - g_2 + 2\lambda w_2$$

Since $g_1 = g_2 = 1$, we know $w_1 - 1 = 0$. Thus, $w_1 = 1$.
The derivative of the Lagrange function with respect to λ is

$$\frac{\partial \mathcal{L}}{\partial \lambda} = ([0\ 1]\vec{w})^2 - 4$$

By setting these to 0, we can solve for $\hat{w}$ and λ.

$$0 = ([0\ 1]\hat{w})^2 - 4$$
$$\sqrt{4} = ([0\ 1]\hat{w})$$
$$\hat{w}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \hat{w}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Substitute these into the partial derivative of the Lagrange function with respect to w_2 to find λ and with $g = [1; 1]$. First, using $\hat{w}_1$

$$0 = 2 - 1 + 2\lambda_1(2)$$
$$\lambda_1 = -0.25$$

Then, using $\hat{w}_2$

$$0 = -2 - 1 + 2\lambda_2(-2)$$
$$\lambda_2 = -0.75$$

Based on the dual feasibility condition, assuming that the constraint is active resulted in invalid solution. Let's suppose that the constraint is inactive and that $\lambda = 0$. In that case, it still holds that $w_1 = 1$. Setting the partial derivative of the Lagrange function with respect to w_2 to 0 means

$$w_2 = g_2 = 1$$

Verify that the new solution satisfies the KKT conditions. Starting with primal feasibility:

$$\begin{aligned} A\hat{w} &\leq 4 \\ 1 &\leq 4 \end{aligned}$$

Dual feasibility:

$$\begin{aligned} \lambda &\geq 0 \\ 0 &\geq 0 \end{aligned}$$

Stationary:

$$\begin{aligned} \begin{bmatrix} \hat{w}_1 - g_1 \\ \hat{w}_2 - g_2 + 2\lambda\hat{w}_2 \end{bmatrix} &= \vec{0} \\ \begin{bmatrix} 1 - 1 \\ 1 - 1 - 0 \end{bmatrix} &= \vec{0} \end{aligned}$$

Complementary slackness:

$$\begin{aligned} \lambda(A\hat{w} - b) &= 0 \\ 0(1 - 4) &= 0 \end{aligned}$$

All conditions have been met. Therefore, $\hat{w} = [1; 1]$ and $\hat{\lambda} = 0$. Since $\hat{\lambda}$ is zero, we can say that the constraint is inactive and the minimizer is not on the boundary of the inequality constraint.

(c) Substituting the property function into the Lagrange formula, we have

$$\mathcal{L}(\vec{w}, \lambda) = \frac{1}{2}[\vec{w} - \vec{g}]^T[\vec{w} - \vec{g}] + \lambda(\vec{w}^T\vec{w} - 1)$$

The derivative of the Lagrange function with respect to $\vec{w}$ is

$$\frac{\partial \mathcal{L}}{\partial \vec{w}} = \vec{w} - \vec{g} + 2\lambda\vec{w}$$

Set this equal to the zero vector and rearrange

$$\begin{aligned} \vec{0} &= \vec{w} - \vec{g} + 2\lambda\vec{w} \\ \vec{g} &= \vec{w} + 2\lambda\vec{w} \\ \vec{g} &= \vec{w}(2\lambda + 1) \\ \vec{w} &= \frac{\vec{g}}{2\lambda + 1} \end{aligned}$$

The derivative of the Lagrange function with respect to λ is

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \vec{w}^T\vec{w} - 1$$

By setting the above to 0, we can solve for λ in terms of $\vec{g}$ by substituting the value of $\vec{w}$ in terms of λ

$$\begin{aligned}
0 &= \vec{w}^T\vec{w} - 1 \\
0 &= \frac{\vec{g}^T}{2\lambda+1}[\frac{\vec{g}}{2\lambda+1}] - 1 \\
1 &= \frac{\vec{g}^T\vec{g}}{(2\lambda+1)^2} \\
0 &= 4\lambda^2 + 4\lambda + 1 - \vec{g}^T\vec{g}
\end{aligned}$$

Consider if $\vec{g}^T\vec{g} \leq 1$. Let the constant term in the quadratic be $c_1 = 1 - \vec{g}^T\vec{g}$ where $\vec{g}^T\vec{g} \leq 1$

$$\begin{aligned}
0 &= 4\lambda^2 + 4\lambda + c_1 \\
\lambda_p &= \frac{-1+\sqrt{c_1+1}}{2}, \lambda_q = \frac{-1-\sqrt{c_1+1}}{2}
\end{aligned}$$

Substitute these into the $\vec{w}$ expressed in terms of λ

$$\begin{aligned}
\vec{w} &= \frac{\vec{g}}{2\lambda+1} \\
\hat{w}_p &= \frac{\vec{g}}{\sqrt{c_1+1}}, \hat{w}_q = \frac{\vec{g}}{-\sqrt{c_1+1}}
\end{aligned}$$

Based on the dual feasibility condition, λ_q and $\vec{w}_q$ cannot be valid solutions since $-\sqrt{c_1+1}$ will always be negative. Verify that λ_p and $\hat{w}_p$ with $p_c(\vec{g}) \leq 1$ are valid solutions using the KKT conditions. Primal feasibility:

$$\begin{aligned}
A\hat{w}_p &\leq 1 \\
\frac{\vec{g}^T}{\sqrt{c_1+1}}(\frac{\vec{g}}{\sqrt{c_1+1}}) &\leq 1 \\
\frac{p_c(\vec{g})}{c_1+1} &\leq 1
\end{aligned}$$

The primal feasibility inequality above is true because $p_c(\vec{g}) \leq 1$. Dual feasibility:

$$\begin{aligned}
\lambda_p &\geq 0 \\
\frac{-1+\sqrt{c_1+1}}{2} &\geq 0
\end{aligned}$$

The dual feasibility inequality above is true because c_1 is at least 0 meaning λ is at least 0. Stationarity:

$$\begin{aligned}
\vec{0} &= \hat{w}_p - \vec{g} + 2\lambda_p\hat{w}_p \\
\vec{0} &= \frac{\vec{g}}{\sqrt{c_1+1}} - \vec{g} + 2\frac{-1+\sqrt{c_1+1}}{2}\frac{\vec{g}}{\sqrt{c_1+1}} \\
\vec{0} &= \frac{\vec{g}}{\sqrt{c_1+1}} - \vec{g} + (-1+\sqrt{c_1+1})\frac{\vec{g}}{\sqrt{c_1+1}} \\
\vec{0} &= \frac{\vec{g}}{\sqrt{c_1+1}}(\sqrt{c_1+1}) - \vec{g} \\
\vec{0} &= \vec{g} - \vec{g}
\end{aligned}$$

Complementary slackness:

$$\lambda_p(A\hat{w}_p - b) = 0$$
$$\frac{-1+\sqrt{c_1+1}}{2}\left(\frac{\vec{g}^T}{\sqrt{c_1+1}}\left(\frac{\vec{g}}{\sqrt{c_1+1}}\right) - 1\right) = 0$$
$$(-1+\sqrt{c_1+1})\left(\frac{1-c_1}{c_1+1} - 1\right) = 0$$
$$(-1+\sqrt{c_1+1}) = 0$$
$$c_1 = 0$$
$$\text{Or}$$
$$\left(\frac{1-c_1}{c_1+1} - 1\right) = 0$$
$$c_1 = 0$$

This means that to satisfy the complementary slackness equality, $p_c(\vec{g}) = 1$. This would still be a viable solution since the condition for the hyperparameter was that $p_c(\vec{g}) \leq 1$. The expressions for λ_p and $\hat{w}_p$ can further be simplified since the value of $p_c(\vec{g})$ has been determined. This will be done later.

Consider if $\vec{g}^T\vec{g} > 1$. Let the constant term in the quadratic be $c_2 = 1 - \vec{g}^T\vec{g}$ where $\vec{g}^T\vec{g} > 1$

$$0 = 4\lambda^2 + 4\lambda + c_2$$
$$\lambda_r = \frac{-1+\sqrt{c_2+1}}{2}, \lambda_s = \frac{-1-\sqrt{c_2+1}}{2}$$

Substitute these into the $\vec{w}$ expressed in terms of λ

$$\vec{w} = \frac{\vec{g}}{2\lambda+1}$$
$$\hat{w}_r = \frac{\vec{g}}{\sqrt{c_2+1}}, \hat{w}_s = \frac{\vec{g}}{-\sqrt{c_2+1}}$$

Based on the dual feasibility condition, λ_s and $\vec{w}_s$ cannot be valid solutions since $-\sqrt{c_2+1}$ will always be negative or complex. Verify that λ_r and $\hat{w}_r$ with $p_c(\vec{g}) > 1$ are valid solutions using the KKT conditions. Primal feasibility:

$$A\hat{w}_p \leq 1$$
$$\frac{\vec{g}^T}{\sqrt{c_2+1}}\left(\frac{\vec{g}}{\sqrt{c_2+1}}\right) \leq 1$$
$$\frac{p_c(\vec{g})}{2-p_c(\vec{g})} \leq 1$$

The primal feasibility inequality above is true if and only if $p_c(\vec{g}) > 2$. This is acceptable for now, but note this finding for the other conditions. Dual feasibility:

$$\lambda_r \geq 0$$
$$\frac{-1+\sqrt{c_2+1}}{2} \geq 0$$

The dual feasibility inequality above is true if and only if $c_2 \geq 0$ meaning $p_c(\vec{g}) \leq 1$. This contradicts the original condition that $p_c(\vec{g}) > 1$ and what was found when validating dual feasibility. Thus, for $p_c(\vec{g}) > 1$, there are no valid solutions that satisfy the KKT conditions.

Therefore, for this constrained optimization problem, there is one valid solution of $\hat{w} = \vec{g}$ and $\hat{\lambda} = 0$ where the hyperparameter $\vec{g}$ must satisfy $p_c(\vec{g}) = 1$. Since $\lambda = 0$, we can say that this constraint is inactive.