CS1.404: Assignment 3

Himanshu Singh

May 6, 2024

KKT Condition

For the convex optimization problem,

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$$
$$\text{s.t. } h_j(\mathbf{x}) \le 0 \qquad \forall j \in \{1,2,\dots,l\}$$

$\mathbf{x}^*$ is an optimal solution, if there exists multipliers $\{\lambda_j\}_{j=1}^{l}$, such that the following equations hold.

$$\begin{aligned}
\nabla f(\mathbf{x}^*) + \sum_{j=1}^{l} \lambda_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0} \\
\lambda_j h_j(\mathbf{x}^*) &= 0 \qquad \forall j \in \{1,2,\dots,l\} \\
\lambda_j &\ge 0 \qquad \forall j \in \{1,2,\dots,l\}
\end{aligned} \tag{1}$$

1 Trid Function

$$f(\mathbf{x}) = \sum_{i=1}^{d} (x_i - 1)^2 - \sum_{i=2}^{d} x_{i-1}x_i$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 - x_2 - 2 \\ 2x_2 - x_1 - x_3 - 2 \\ 2x_3 - x_2 - x_4 - 2 \\ \vdots \\ 2x_{d-1} - x_{d-2} - x_d - 2 \\ 2x_d - x_{d-1} - 2 \end{bmatrix}$$

1.1 Test Case 0

$$h_1(\mathbf{x}) = x_1^2 - 2x_2$$
$$\nabla h_1(\mathbf{x}) = \begin{bmatrix} 2x_1 & -2 \end{bmatrix}^T$$

Substituting in (1), we get

$$2x_1 - x_2 - 2 + 2\lambda_1 x_1 = 0$$

$$
\begin{aligned}
2x_2 - x_1 - 2 - 2\lambda_1 &= 0 \\
\lambda_1(x_1^2 - 2x_2) &= 0
\end{aligned} \tag{2}
$$

Case 1 When $\lambda_1 = 0$, the system (2) reduces to

$$
\begin{aligned}
2x_1 - x_2 - 2 &= 0 \\
2x_2 - x_1 - 2 &= 0
\end{aligned} \tag{4}
$$

Solving (4), we get $x_1 = x_2 = 2$.

Case 2 When $\lambda_1 \neq 0$, the system (2) reduces to

$$
\begin{aligned}
2x_1 - x_2 - 2 + 2\lambda_1 x_1 &= 0 \\
2x_2 - x_1 - 2 - 2\lambda_1 &= 0 \\
x_1^2 - 2x_2 &= 0
\end{aligned} \tag{6}
$$

Eliminating x_2 from (6), we get

$$
\begin{aligned}
2x_1 - \frac{x_1^2}{2} - 2 + 2\lambda_1 x_1 &= 0 \\
x_1^2 - x_1 - 2 - 2\lambda_1 &= 0
\end{aligned} \tag{7}
$$

Eliminating λ_1 from (7), we get

$$
x_1^3 - \frac{3x_1^2}{2} - 2 = 0 \tag{8}
$$

Solving (8) we get $x_1 = 2$ as the only real solution. Substituting back in (7), we get $\lambda_1 = 0$, contradicting our assumption that $\lambda_1 \neq 0$.
We thus conclude that $(\mathbf{x}^*, \lambda^*) = \left(\begin{bmatrix} 2 & 2 \end{bmatrix}^T, 0 \right)$ is the only real solution to (2).

1.2 Test Case 1

$$
\begin{aligned}
h_1(\mathbf{x}) &= x_1^2 - x_2^2 + 1 \\
\nabla h_1(\mathbf{x}) &= \begin{bmatrix} 2x_1 & -2x_2 \end{bmatrix}^T
\end{aligned}
$$

Substituting in (1), we get

$$
\begin{aligned}
2x_1 - x_2 - 2 + 2\lambda_1 x_1 &= 0 \\
2x_2 - x_1 - 2 - 2\lambda_1 x_2 &= 0 \\
\lambda_1(x_1^2 - x_2^2 + 1) &= 0
\end{aligned} \tag{9}
$$

Case 1 When $\lambda_1 = 0$, the system (9) reduces to

$$
\begin{aligned}
2x_1 - x_2 - 2 &= 0 \\
2x_2 - x_1 - 2 &= 0
\end{aligned} \tag{11}
$$

Solving (11), we get $x_1 = x_2 = 2$.

Case 2 When $\lambda_1 \neq 0$, the system (9) reduces to

$$\begin{aligned} 2x_1 - x_2 - 2 + 2\lambda_1 x_1 &= 0 \\ 2x_2 - x_1 - 2 - 2\lambda_1 x_2 &= 0 \\ x_1^2 - x_2^2 + 1 &= 0 \end{aligned} \tag{13}$$

Eliminating x_2 from (13), we get

$$\begin{aligned} 2x_1 \mp \sqrt{x_1^2+1} - 2 + 2\lambda_1 x_1 &= 0 \\ \pm 2\sqrt{x_1^2+1} - x_1 - 2 \mp 2\lambda_1 \sqrt{x_1^2+1} &= 0 \end{aligned} \tag{14}$$

Eliminating λ_1 from (14), we get

$$\mp \frac{x_1^2 + 2x_1}{\sqrt{x_1^2+1}} + 4x_1 \mp \sqrt{x_1^2+1} - 2 = 0 \tag{15}$$

Solving (15) we get $x_1 = \{1.8991, 0.16094\}$ as the real solutions. Substituting back in (14), we get $\lambda_1 = \{0.90167, 2.06674\}$ respectively. Substituting back in (13), we get $x_2 = \{2.14629, -1.01287\}$ respectively.
We thus conclude that $(\mathbf{x}^*, \lambda^*) = \left\{ \left(\begin{bmatrix} 2 & 2 \end{bmatrix}^T, 0 \right), \left(\begin{bmatrix} 1.8991 & 2.14629 \end{bmatrix}^T, 0.09167 \right), \left(\begin{bmatrix} 0.16094 & -1.01287 \end{bmatrix}^T, 2.0667\right.\right.$
as the real solutions to (9).

1.3 Test Case 2

$$\begin{aligned} h_1(\mathbf{x}) &= -1 - x_1 \\ h_2(\mathbf{x}) &= x_1 - 1 \\ h_3(\mathbf{x}) &= -1 - x_2 \\ h_4(\mathbf{x}) &= x_2 - 1 \\ \nabla h_1(\mathbf{x}) &= \begin{bmatrix} -1 & 0 \end{bmatrix}^T \\ \nabla h_2(\mathbf{x}) &= \begin{bmatrix} 1 & 0 \end{bmatrix}^T \\ \nabla h_3(\mathbf{x}) &= \begin{bmatrix} 0 & -1 \end{bmatrix}^T \\ \nabla h_4(\mathbf{x}) &= \begin{bmatrix} 0 & 1 \end{bmatrix}^T \end{aligned}$$

Substituting in (1), we get

$$\begin{aligned} 2x_1 - x_2 - 2 - \lambda_1 + \lambda_2 &= 0 \\ 2x_2 - x_1 - 2 - \lambda_3 + \lambda_4 &= 0 \\ \lambda_1(x_1 + 1) &= 0 \\ \lambda_2(x_1 - 1) &= 0 \\ \lambda_3(x_2 + 1) &= 0 \\ \lambda_4(x_2 - 1) &= 0 \end{aligned} \tag{16}$$

We observe that when $\mathbf{x} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, the system (16) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}^T$.

1.4 Test Case 3

$$
\begin{aligned}
h_1(\mathbf{x}) &= -x_1 \\
h_2(\mathbf{x}) &= x_1 - 3 \\
h_3(\mathbf{x}) &= -x_2 \\
h_4(\mathbf{x}) &= x_2 - 3 \\
\nabla h_1(\mathbf{x}) &= \begin{bmatrix} -1 & 0 \end{bmatrix}^T \\
\nabla h_2(\mathbf{x}) &= \begin{bmatrix} 1 & 0 \end{bmatrix}^T \\
\nabla h_3(\mathbf{x}) &= \begin{bmatrix} 0 & -1 \end{bmatrix}^T \\
\nabla h_4(\mathbf{x}) &= \begin{bmatrix} 0 & 1 \end{bmatrix}^T
\end{aligned}
$$

Substituting in (1), we get

$$
\begin{aligned}
2x_1 - x_2 - 2 - \lambda_1 + \lambda_2 &= 0 \\
2x_2 - x_1 - 2 - \lambda_3 + \lambda_4 &= 0 \\
\lambda_1 x_1 &= 0 \\
\lambda_2 (x_1 - 3) &= 0 \\
\lambda_3 x_2 &= 0 \\
\lambda_4 (x_2 - 3) &= 0 \qquad (17)
\end{aligned}
$$

We observe that when $\mathbf{x} = \begin{bmatrix} 2 & 2 \end{bmatrix}^T$, the system (17) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$.

1.5 Test Case 4

$$
\begin{aligned}
h_1(\mathbf{x}) &= 3 - x_1 \\
h_2(\mathbf{x}) &= x_1 - 4 \\
h_3(\mathbf{x}) &= 3 - x_2 \\
h_4(\mathbf{x}) &= x_2 - 4 \\
\nabla h_1(\mathbf{x}) &= \begin{bmatrix} -1 & 0 \end{bmatrix}^T \\
\nabla h_2(\mathbf{x}) &= \begin{bmatrix} 1 & 0 \end{bmatrix}^T \\
\nabla h_3(\mathbf{x}) &= \begin{bmatrix} 0 & -1 \end{bmatrix}^T \\
\nabla h_4(\mathbf{x}) &= \begin{bmatrix} 0 & 1 \end{bmatrix}^T
\end{aligned}
$$

Substituting in (1), we get

$$
\begin{aligned}
2x_1 - x_2 - 2 - \lambda_1 + \lambda_2 &= 0 \\
2x_2 - x_1 - 2 - \lambda_3 + \lambda_4 &= 0 \\
\lambda_1 (x_1 - 3) &= 0
\end{aligned}
$$

$$\begin{aligned}
\lambda_2(x_1-4) &= 0 \\
\lambda_3(x_2-3) &= 0 \\
\lambda_4(x_2-4) &= 0
\end{aligned} \tag{18}$$

We observe that when $\mathbf{x} = \begin{bmatrix}3 & 3\end{bmatrix}^T$, the system (18) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix}1 & 0 & 1 & 0\end{bmatrix}^T$.

2 Matyas Function

$$f(\mathbf{x}) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix}0.52x_1 - 0.48x_2 \\ 0.52x_2 - 0.48x_1\end{bmatrix}$$

2.1 Test Case 5

$$\begin{aligned}
h_1(\mathbf{x}) &= -x_1 \\
h_2(\mathbf{x}) &= x_1 - 1 \\
h_3(\mathbf{x}) &= -x_2 \\
h_4(\mathbf{x}) &= x_2 - 1 \\
\nabla h_1(\mathbf{x}) &= \begin{bmatrix}-1 & 0\end{bmatrix}^T \\
\nabla h_2(\mathbf{x}) &= \begin{bmatrix}1 & 0\end{bmatrix}^T \\
\nabla h_3(\mathbf{x}) &= \begin{bmatrix}0 & -1\end{bmatrix}^T \\
\nabla h_4(\mathbf{x}) &= \begin{bmatrix}0 & 1\end{bmatrix}^T
\end{aligned}$$

Substituting in (1), we get

$$\begin{aligned}
0.52x_1 - 0.48x_2 - \lambda_1 + \lambda_2 &= 0 \\
0.52x_2 - 0.48x_1 - \lambda_3 + \lambda_4 &= 0 \\
\lambda_1 x_1 &= 0 \\
\lambda_2(x_1-1) &= 0 \\
\lambda_3 x_2 &= 0 \\
\lambda_4(x_2-1) &= 0
\end{aligned} \tag{19}$$

We observe that when $\mathbf{x} = \begin{bmatrix}0 & 0\end{bmatrix}^T$, the system (19) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix}^T$.

2.2 Test Case 6

$$\begin{aligned}
h_1(\mathbf{x}) &= 1 - x_1 \\
h_2(\mathbf{x}) &= x_1 - 2 \\
h_3(\mathbf{x}) &= 1 - x_2 \\
h_4(\mathbf{x}) &= x_2 - 2
\end{aligned}$$

$$
\begin{aligned}
\nabla h_1(\mathbf{x}) &= \begin{bmatrix}-1 & 0\end{bmatrix}^T \\
\nabla h_2(\mathbf{x}) &= \begin{bmatrix}1 & 0\end{bmatrix}^T \\
\nabla h_3(\mathbf{x}) &= \begin{bmatrix}0 & -1\end{bmatrix}^T \\
\nabla h_4(\mathbf{x}) &= \begin{bmatrix}0 & 1\end{bmatrix}^T
\end{aligned}
$$

Substituting in (1), we get

$$
\begin{aligned}
0.52x_1 - 0.48x_2 - \lambda_1 + \lambda_2 &= 0 \\
0.52x_2 - 0.48x_1 - \lambda_3 + \lambda_4 &= 0 \\
\lambda_1(x_1 - 1) &= 0 \\
\lambda_2(x_1 - 2) &= 0 \\
\lambda_3(x_2 - 1) &= 0 \\
\lambda_4(x_2 - 2) &= 0
\end{aligned} \tag{20}
$$

We observe that when $\mathbf{x} = \begin{bmatrix}1 & 1\end{bmatrix}^T$, the system (20) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix}0.04 & 0 & 0.04 & 0\end{bmatrix}^T$.

2.3 Test Case 7

$$
\begin{aligned}
h_1(\mathbf{x}) &= -1 - x_1 \\
h_2(\mathbf{x}) &= x_1 + 0.5 \\
h_3(\mathbf{x}) &= -0.5 - x_2 \\
h_4(\mathbf{x}) &= x_2 - 0.5 \\
\nabla h_1(\mathbf{x}) &= \begin{bmatrix}-1 & 0\end{bmatrix}^T \\
\nabla h_2(\mathbf{x}) &= \begin{bmatrix}1 & 0\end{bmatrix}^T \\
\nabla h_3(\mathbf{x}) &= \begin{bmatrix}0 & -1\end{bmatrix}^T \\
\nabla h_4(\mathbf{x}) &= \begin{bmatrix}0 & 1\end{bmatrix}^T
\end{aligned}
$$

Substituting in (1), we get

$$
\begin{aligned}
0.52x_1 - 0.48x_2 - \lambda_1 + \lambda_2 &= 0 \\
0.52x_2 - 0.48x_1 - \lambda_3 + \lambda_4 &= 0 \\
\lambda_1(x_1 + 1) &= 0 \\
\lambda_2(x_1 + 0.5) &= 0 \\
\lambda_3(x_2 + 0.5) &= 0 \\
\lambda_4(x_2 - 0.5) &= 0
\end{aligned} \tag{21}
$$

We observe that when $\mathbf{x} = \begin{bmatrix}-0.5 & -0.4615\end{bmatrix}^T$, the system (21) is satisifed for $\boldsymbol{\lambda} = \begin{bmatrix}0 & 0.03848 & 0 & 0\end{bmatrix}^T$.