Lecture: KKT conditions for NLP with inequality constraints

1. KKT conditions for general nonlinear optimization problems with inequality constraints.

General nonlinear optimization problems with inequality constraints

Consider

$$(NLP_{\leq}) \quad \left[\begin{array}{rl} \text{minimize} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \leq 0, \;\; i = 1, \ldots, m \end{array} \right]$$

where $f(\mathbf{x})$ and $g_i(\mathbf{x})$ are real valued functions.

The feasible region is given by

$$\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0, \;\; i = 1, \ldots, m\}$$

We want to derive necessary optimality conditions.

Definition 1. *For* $\mathbf{x} \in \mathcal{F}$ *we let* $\mathcal{I}_a(\mathbf{x})$ *denote the* index set for active constraints *in the point* $\mathbf{x}$, *i.e.*, $\mathcal{I}_a(\mathbf{x}) = \{i \in \{1, \ldots, m\} : g_i(\mathbf{x}) = 0\}$.

Regularity for $(NLP_{\leq})$

We will look at optimality conditions that will hold in all points except those that are not regular, so we want to have as few points as possible that are not regular.

First attempt:

Definition 2. *A feasible solution* $\mathbf{x} \in \mathcal{F}$ *is a regular point to* $(NLP_{\leq})$ *if* $\nabla g_i(\mathbf{x})$ *for* $i \in \mathcal{I}_a(\mathbf{x})$ *are linearly independent.*

With a stronger condition we can get optimality conditions that are applicable for more problems.

Regularity for $(NLP_{\leq})$

The regularity condition in Definition 2 can be replaced with the stronger condition:

Definition 3. *A feasible solution* $\mathbf{x} \in \mathcal{F}$ *with* $\mathcal{I}_a(\mathbf{x})$ *non-empty is a regular point to* $(NLP_{\leq})$ *if there does not exist scalars* $v_i \geq 0$, $i \in \mathcal{I}_a(\mathbf{x})$, *such that*

$$\sum_{i \in \mathcal{I}_a(\mathbf{x})} v_i > 0$$

and

$$\sum_{i \in \mathcal{I}_a(\mathbf{x})} v_i \nabla g_i(\mathbf{x}) = 0.$$

A feasible point with $\mathcal{I}_a(\mathbf{x})$ *empty is always a regular point.*

Theorem 1 (KKT for general problems with inequality constraints)**.**
Assume that $\hat{\mathbf{x}}$ *is a regular point to* $(NLP_{\leq})$ *and a local optimal solution.*
Then there exists a vector $\hat{\mathbf{y}} \in \mathbf{R}^m$ *such that*

(1) $\nabla f(\hat{\mathbf{x}}) + \hat{\mathbf{y}}^T \nabla \mathbf{g}(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$

(2) $\mathbf{g}(\hat{\mathbf{x}}) \leq 0$,

(3) $\hat{\mathbf{y}} \geq 0$,

(4) $\hat{\mathbf{y}}^T \mathbf{g}(\hat{\mathbf{x}}) = 0$.

The conditions (1) – (4) can be made more explicit. We have that

$$\hat{\mathbf{y}}^{\mathsf{T}}\mathbf{g}(\hat{\mathbf{x}}) = \sum_{i=1}^{m} \hat{y}_i g_i(\hat{\mathbf{x}}) = 0$$

Since $g_i(\hat{\mathbf{x}}) \leq 0$ and $\hat{y}_i \geq 0$ it follows that $\hat{y}_i g_i(\hat{\mathbf{x}}) = 0$, $i = 1, \ldots, m$.
We then get the equivalent conditions

(2′) $g_i(\hat{\mathbf{x}}) \leq 0$, $i = 1, \ldots, m$,

(3′) $\hat{y}_i \geq 0$, $i = 1, \ldots, m$,

(4′) $\hat{y}_i \cdot g_i(\hat{\mathbf{x}}) = 0$, $i = 1, \ldots, m$.

Theorem 2 (KKT for general problems with inequality constraints). *Assume that* $\hat{\mathbf{x}}$ *is a regular point to* $(NLP_{\leq})$ *and a local optimal solution.*

Then there exists a vector $\hat{\mathbf{y}} \in \mathbf{R}^m$ *such that*

(1′) $\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{m} \hat{\mathbf{y}}_i \nabla g_i(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$

(2′) $g_i(\hat{\mathbf{x}}) \leq 0$, $i = 1, \ldots, m$,

(3′) $\hat{\mathbf{y}}_i \geq 0$, $i = 1, \ldots, m$,

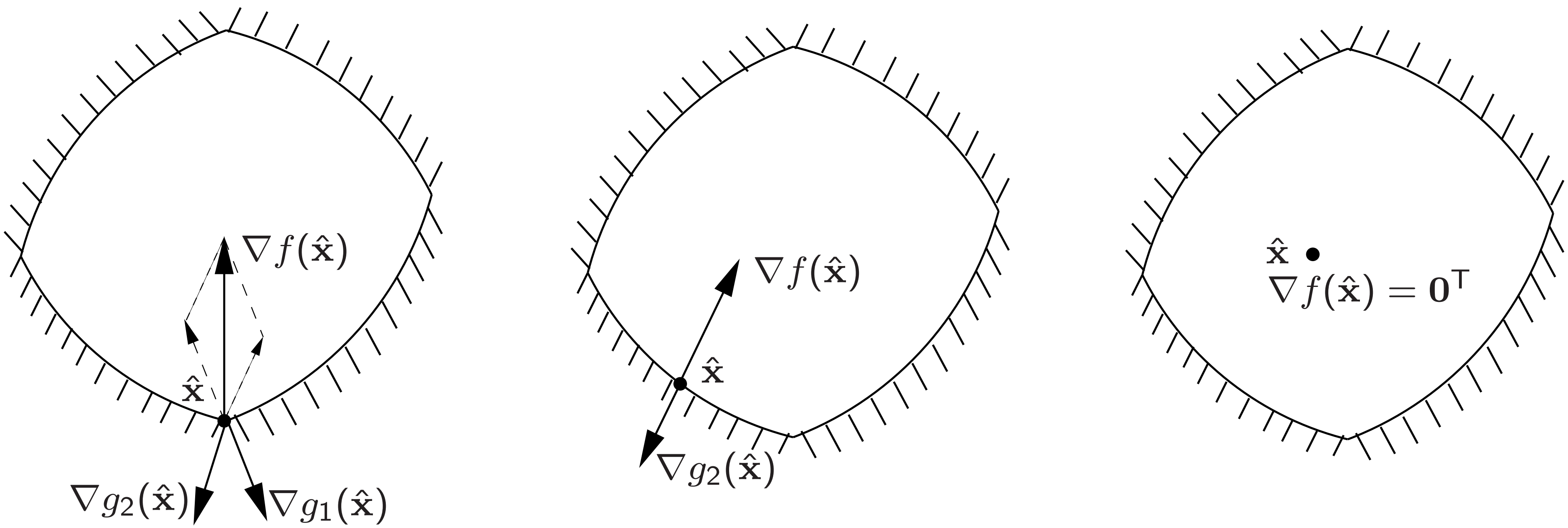
(4′) $\hat{\mathbf{y}}_i \cdot g_i(\hat{\mathbf{x}}) = 0$, $i = 1, \ldots, m$.

Geometric interpretation

The complementarity condition (4′) implies that if $g_i(\hat{\mathbf{x}}) < 0$ then $y_i = 0$. Therefore, condition (1′) can be written

$$\nabla f(\hat{\mathbf{x}}) = - \sum_{i:g_i(\hat{\mathbf{x}})=0} \hat{y}_i \nabla g_i(\hat{\mathbf{x}})$$

this means that the gradient is a negative linear combination of the gradients of the binding (active) constraints.



An example

$$\begin{aligned} \text{minimize} \quad & (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{s.t.} \quad & 2x_1 + x_2 - 6 \le 0, \\ & x_1 + 2x_2 - 6 \le 0 \end{aligned}$$

Here $f(\mathbf{x}) = (x_1 - 3)^2 + (x_2 - 2)^2$,
$g_1(\mathbf{x}) = 2x_1 + x_2 - 6$ and $g_2(\mathbf{x}) = x_1 + 2x_2 - 6$, then

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 3) & 2(x_2 - 2) \end{bmatrix},$$

$$\nabla g_1(\mathbf{x}) = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad \nabla g_2(\mathbf{x}) = \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

The gradients of g are linearly independent so all points are regular.

Solving the KKT-conditions

We can determine all solutions to the KKT-conditions by considering all combinations of active and non-active constraints.
(practical only on small problems)

Four cases:

1. No constraints active $\mathcal{I}_a(\mathbf{x}) = \emptyset$,
   i.e. $g_1(\mathbf{x}) < 0$ and $g_2(\mathbf{x}) < 0$.

2. First active and second not-active $\mathcal{I}_a(\mathbf{x}) = \{1\}$,
   i.e. $g_1(\mathbf{x}) = 0$ and $g_2(\mathbf{x}) < 0$.

3. First not-active and second active $\mathcal{I}_a(\mathbf{x}) = \{2\}$,
   i.e. $g_1(\mathbf{x}) < 0$ and $g_2(\mathbf{x}) = 0$.

4. Both constraints active $\mathcal{I}_a(\mathbf{x}) = \{1, 2\}$,
   i.e. $g_1(\mathbf{x}) = 0$ and $g_2(\mathbf{x}) = 0$.

Case 1: $\mathcal{I}_a(\mathbf{x}) = \emptyset$

From KKT(4) we get that both $y_1 = 0$ and $y_2 = 0$.

Then, KKT(1) is

$$\nabla f(\mathbf{x}) = \left[\begin{array}{cc} 2(x_1 - 3) & 2(x_2 - 2) \end{array} \right] = 0,$$

which implies that $x_1 = 3$ and $x_2 = 2$.

But since KKT(2) is not satisifed, $2x_1 + x_2 - 6 = 2 \not\leq 0$, (i.e. $\mathbf{x}$ is not feasible) for these values of $\mathbf{x}$, it can not be a local minimum.

Case 2: $\mathcal{I}_a(\mathbf{x}) = \{1\}$

From KKT(4) we get that $y_2 = 0$. Then KKT(1) is

$$\nabla f(\mathbf{x}) + y_1 \nabla g_1(\mathbf{x}) = \left[\begin{array}{cc} 2(x_1 - 3) & 2(x_2 - 2) \end{array}\right] + y_1 \left[\begin{array}{cc} 2 & 1 \end{array}\right] = 0,$$

which implies that $x_1 = 3 - y_1$ and $x_2 = 2 - y_1/2$.
The assumption $g_1(\mathbf{x}) = 0$ gives now with these x_1 and x_2

$$2(3 - y_1) + (2 - y_1/2) - 6 = 0 \qquad \Rightarrow \qquad y_1 = 4/5$$

Tis y_1 (and y_2) satisfies KKT(3), and then we get
$x_1 = 3 - 4/5 = 11/5$ and $x_2 = 2 - 1/2(4/5) = 8/5$.

Finally, $g_2(\mathbf{x}) = 11/5 + 2(8/5) - 6 = -3/5 \leq 0$ so KKT(2) is satisfied.

All KKT-conditions are satisfied!

Case 3: $\mathcal{I}_a(\mathbf{x}) = \{2\}$

From KKT(4) we get $y_1 = 0$. Then KKT(1) gives that

$$\nabla f(\mathbf{x}) + y_2 \nabla g_2(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 3) & 2(x_2 - 2) \end{bmatrix} + y_2 \begin{bmatrix} 1 & 2 \end{bmatrix} = 0,$$

hence $x_1 = 3 - y_2/2$ and $x_2 = 2 - y_2$.

The assumption $g_2(\mathbf{x}) = 0$ gives now, with these x_1 and x_2

$$(3 - y_2/2) + 2(2 - y_2) - 6 = 0 \qquad \Rightarrow \qquad y_2 = 2/5$$

This y_2 (and y_1) satisfies KKT(3), and then
$x_1 = 3 - 1/5 = 14/5$ and $x_2 = 2 - (2/5) = 8/5$.

Finally, $g_1(\mathbf{x}) = 2(14/5) + 8/5 - 6 = 1/5 > 0$ so KKT(2) is *not* satisfied.

Case 4: $\mathcal{I}_a(\mathbf{x}) = \{1, 2\}$

The assumption $g_1(\mathbf{x}) = 2x_1 + x_2 - 6 = 0$ and $g_2(\mathbf{x}) = x_1 + 2x_2 - 6 = 0$ gives that $x_1 = 2$ and $x_2 = 2$.

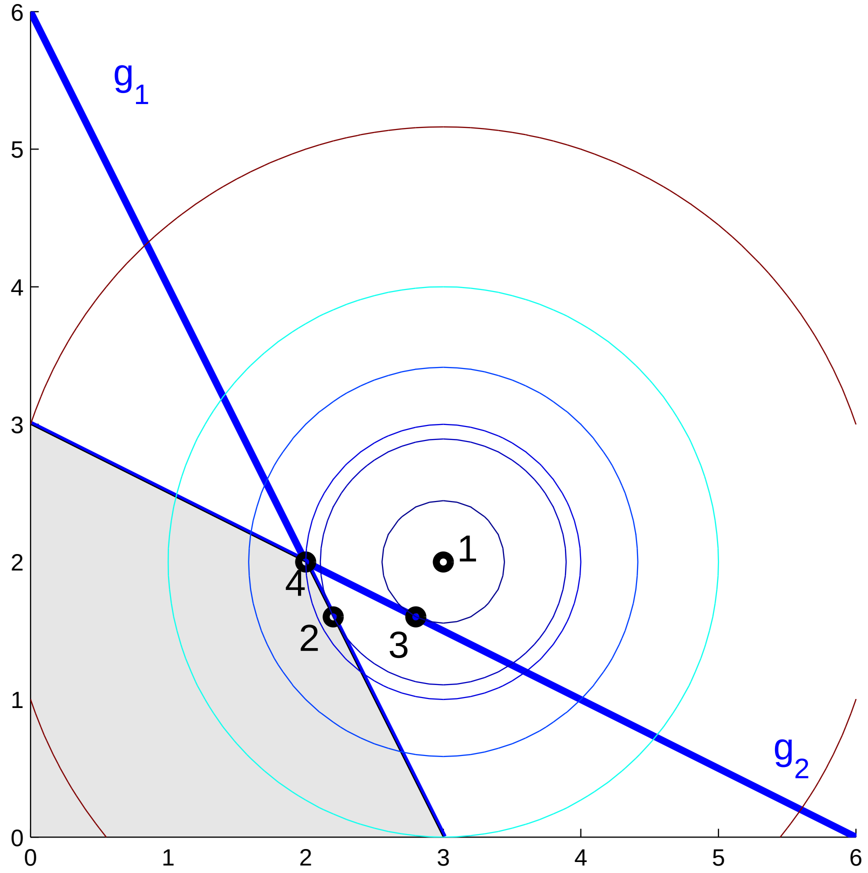
KKT(1) says that $\nabla f(\mathbf{x}) + y_1 \nabla g_1(\mathbf{x}) + y_2 \nabla g_2(\mathbf{x}) = 0$,

$$\begin{bmatrix} 2(x_1 - 3) & 2(x_2 - 2) \end{bmatrix} + y_1 \begin{bmatrix} 2 & 1 \end{bmatrix} + y_2 \begin{bmatrix} 1 & 2 \end{bmatrix} = 0,$$

with x_1 and x_2 inserted giving $-2 + 2y_1 + y_2 = 0$ and $y_1 + 2y_2 = 0$, so $y_1 = -2/3$ and $y_2 = 4/3$.

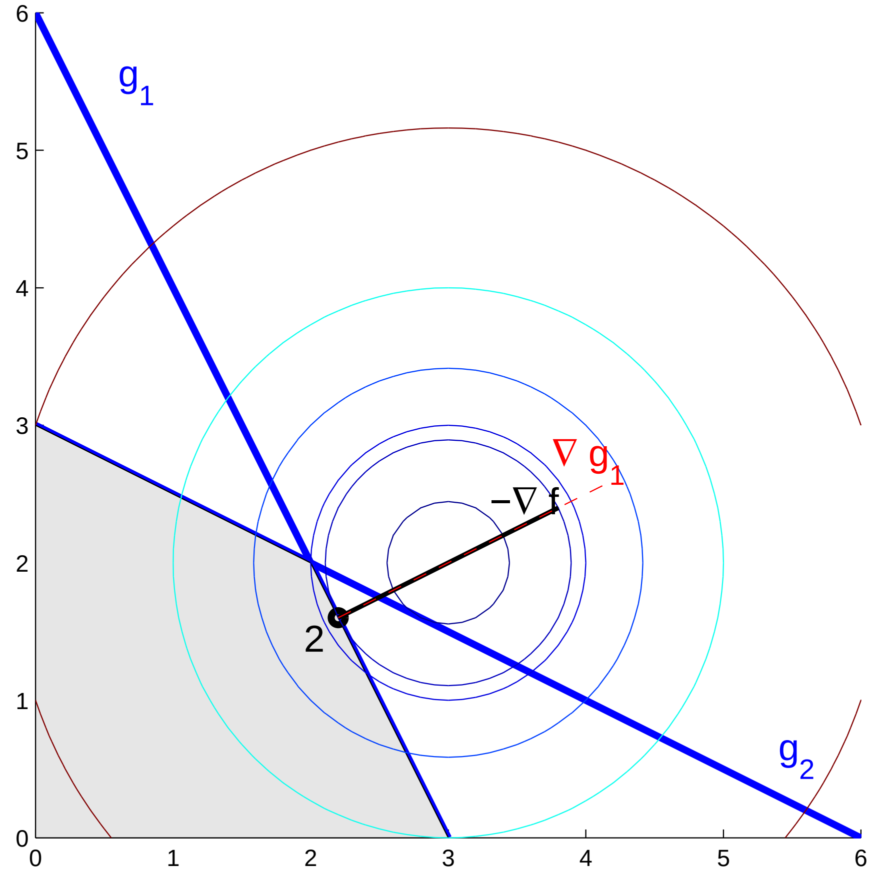
But this $\mathbf{y}$ does not satisfy KKT(3) so it can not be a local minimum.

Graphical illustration



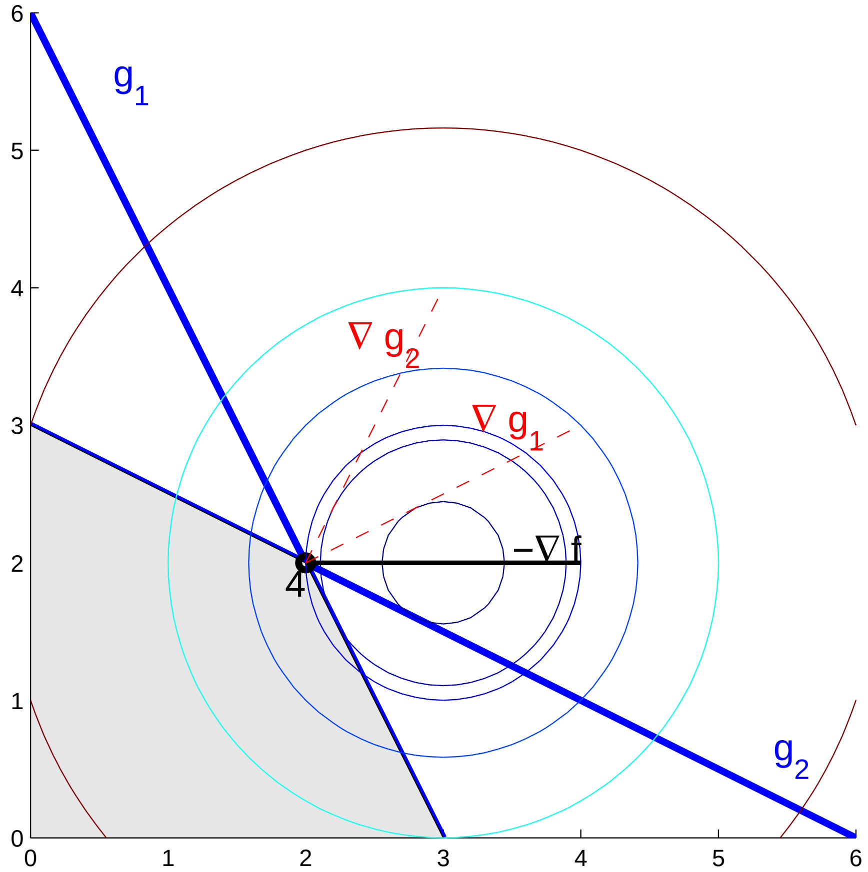
The four “solutions” are depicted in the figure. As we saw above, it is only number 2 that satisfies all the KKT-constraints.

Case 2



Here we can see that minus the gradient to f is a positive linear combination of (the gradient of) the one active constraint.

Case 4



Here we can see that minus the gradient to f is *not* a positive linear combination of the (gradients of the) active constraints.