ECE 68000: MODERN AUTOMATIC CONTROL

Professor Stan Żak

Minimization subject to inequality constraints

Problem statement

- Find a point $\boldsymbol{x} \in \mathbb{R}^N$ that minimizes $f(\boldsymbol{x})$ subject to inequality constraints,

$$\left.\begin{array}{rcl} g_1(\boldsymbol{x}) & \leq & 0 \\ g_2(\boldsymbol{x}) & \leq & 0 \\ & \vdots & \\ g_P(\boldsymbol{x}) & \leq & 0 \end{array}\right\}$$

- Write the above inequality constraints in a compact form as

$$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0},$$

where $\boldsymbol{h} : \mathbb{R}^N \to \mathbb{R}^P$.

- Let $\boldsymbol{x}^*$ be a point satisfying the constraints, that is, $\boldsymbol{g}\left(\boldsymbol{x}^*\right) \leq \boldsymbol{0}$
- Let J be the set of indices j for which $g_j\left(\boldsymbol{x}^*\right) = 0$

Regular point of the constraints

- The point $\boldsymbol{x}^*$ is said to be a regular point of the constraints $\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$ if the gradient vectors,

$$\nabla g_j\left(\boldsymbol{x}^*\right), \quad j \in J,$$

  are linearly independent
- A constraint $g_j(\boldsymbol{x}) \leq 0$ is active at $\boldsymbol{x}^*$ if $g_j\left(\boldsymbol{x}^*\right) = 0$
- The index set J, defined above, contains indices of active constraints

The first-order necessary condition (FONC) for function minimization subject to inequality constraints

Theorem

Let $\boldsymbol{x}^*$ be a regular point and a local minimizer of f subject to $\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$. Then there exists a vector $\boldsymbol{\mu}^* \in \mathbb{R}^P$ such that

1. $\boldsymbol{\mu}^* \geq 0$,
2. $\nabla f\left(\boldsymbol{x}^*\right) + \begin{bmatrix} \nabla g_1\left(\boldsymbol{x}^*\right) & \cdots & \nabla g_P\left(\boldsymbol{x}^*\right) \end{bmatrix} \boldsymbol{\mu}^* = \mathbf{0}$,
3. $\boldsymbol{\mu}^{*\top} \boldsymbol{g}\left(\boldsymbol{x}^*\right) = 0$

Proof of FONC

- Since $\boldsymbol{x}^*$ is a relative minimizer over the constraint set $\{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}$, it is also a minimizer over a subset of the constraint set obtained by setting the active constraints to zero
- Therefore, for the resulting equality constrained problem

$$\nabla f\left(\boldsymbol{x}^*\right) + \begin{bmatrix} \nabla g_1\left(\boldsymbol{x}^*\right) & \cdots & \nabla g_P\left(\boldsymbol{x}^*\right) \end{bmatrix} \boldsymbol{\mu}^* = \mathbf{0},$$

  where $\mu_j^* = 0$ if $g_j\left(\boldsymbol{x}^*\right) < 0$
- This means that $\boldsymbol{\mu}^{*\top} \boldsymbol{g}\left(\boldsymbol{x}^*\right) = 0$
- Thus μ_i may be non-zero only if the corresponding constraint is active, that is, $g_i\left(\boldsymbol{x}^*\right) = 0$

Proving that $\mu^* \geq \mathbf{0}$

- It remains to show that $\mu^* \geq \mathbf{0}$
- By contraposition: suppose that for some $k \in J$, we have $\mu_k^* < 0$
- Consider a surface formed by all other active constraints, that is, the surface

$$\{\boldsymbol{x} : g_j(\boldsymbol{x}) = 0, j \in J, j \neq k\}.$$

- Also consider

$$\nabla g_k\left(\boldsymbol{x}^*\right)^\top \boldsymbol{y} < 0$$

- Let $\boldsymbol{x}(t)$, $t \in [-a, a]$, $a > 0$, be a curve on the surface such that

$$\dot{\boldsymbol{x}}(0) = \boldsymbol{y}$$

- Note that for small t, the curve $\boldsymbol{x}(t)$ is feasible

Minimality of $f\left(\boldsymbol{x}^*\right)$

- Apply the transposition operator to get

$$\nabla f\left(\boldsymbol{x}^*\right)^\top + \sum_{i=1}^{P} \mu_i^* \nabla g_i\left(\boldsymbol{x}^*\right)^\top = \mathbf{0}^\top$$

- Post-multiplying the above by $\boldsymbol{y}$ an taking into account the fact that $\boldsymbol{y}$ belongs to the tangent space to the surface gives

$$\begin{aligned} \nabla f\left(\boldsymbol{x}^*\right)^\top \boldsymbol{y} &= -\mu_k^* \nabla g_k\left(\boldsymbol{x}^*\right)^\top \boldsymbol{y} \\ &< 0 \end{aligned}$$

Finish of the proof of FONC

- Suppose, without loss of generality, that $\|\boldsymbol{y}\|_2 = 1$. Then,

$$\begin{aligned} \frac{df(\boldsymbol{x}(t))}{dt} &= \nabla f\left(\boldsymbol{x}^*\right)^\top \boldsymbol{y} \\ &< 0, \end{aligned}$$

  that is, the rate of increase of f at $\boldsymbol{x}^*$ in the direction $\boldsymbol{y}$ is negative
- This would mean that we could decrease the value of f moving just slightly away from $\boldsymbol{x}^*$ along $\boldsymbol{y}$ while, at the same time, preserving feasibility
- But this contradicts the minimality of $f\left(\boldsymbol{x}^*\right)$
- In sum, if $\boldsymbol{x}^*$ is a relative minimizer then we also have the components of $\boldsymbol{\mu}^*$ all non-negative $\square$
- The vector $\boldsymbol{\mu}^*$ is called the vector of the Karush-Kuhn-Tucker (KKT) multipliers