Lecture 11: KKT Conditions

Shi Pu

School of Data Science (SDS)
The Chinese University of Hong Kong, Shenzhen

Outline

Outline

- The Karush-Kuhn-Tucker conditions are optimality conditions for inequality constrained problems discovered in 1951 (originating from Karush's thesis from 1939).
- Modern nonlinear optimization essentially begins with the discovery of these conditions.

The basic notion that we will require is the one of feasible descent directions.

Definition

Consider the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in C \end{aligned}$$

where f is continuously differentiable over the set $C \subset \mathbb{R}^n$. Then a vector $\mathbf{d} \neq \mathbf{0}$ is called a feasible descent direction at $\mathbf{x} \in C$ if $\nabla f(\mathbf{x})^\top \mathbf{d} < 0$ and there exists $\epsilon > 0$ such that $\mathbf{x} + t\mathbf{d} \in C$ for all $t \in [0, \epsilon]$.

Lemma

Consider the problem

$$(P) \qquad \begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & \mathbf{x} \in C \end{aligned}$$

where f is continuously differentiable over the set C. If $\mathbf{x}^*$ is a local optimal solution of (P), then there are no feasible descent directions at $\mathbf{x}^*$.

Proof.

- By contradiction, assume that there exists a vector $\mathbf{d}$ and $\epsilon_1 > 0$ such that $\mathbf{x} + t\mathbf{d} \in C$ for all $t \in [0, \epsilon_1]$ and $\nabla f(\mathbf{x}^*)^\top \mathbf{d} < 0$.
- By definition of the directional derivative there exists $\epsilon_2 < \epsilon_1$ such that $f(\mathbf{x}^* + t\mathbf{d}) < f(\mathbf{x}^*)$ for all $t \in [0, \epsilon_2] \Rightarrow$ contradiction to the local optimality of $\mathbf{x}^*$.

$\square$

Lemma

Let $\mathbf{x}^*$ be a local minimum of the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \cdots, m \end{aligned}$$

where $f, g_1, \cdots, g_m$ are continuously differentiable over $\mathbb{R}^n$. Let $I(\mathbf{x}^*)$ be the set of active constraints at $\mathbf{x}^*$:

$$I(\mathbf{x}^*) = \{i : g_i(\mathbf{x}^*) = 0\}$$

Then there does not exist a vector $\mathbf{d} \in \mathbb{R}^n$ such that

$$\begin{aligned} &\nabla f(\mathbf{x}^*)^\top \mathbf{d} < 0 \\ &\nabla g_i(\mathbf{x}^*)^\top \mathbf{d} < 0, \ i \in I(\mathbf{x}^*) \end{aligned}$$

Proof.

- Suppose that $\mathbf{d}$ satisfies the system of inequalities.
- Then $\exists\ \epsilon_1 > 0$ such that $f(\mathbf{x}^* + t\mathbf{d}) < f(\mathbf{x}^*)$ and $g_i(\mathbf{x}^* + t\mathbf{d}) < g_i(\mathbf{x}^*) = 0$ for any $t \in (0, \epsilon_1)$ and $i \in I(\mathbf{x}^*)$.
- For any $i \notin I(\mathbf{x}^*)$ we have that $g_i(\mathbf{x}^*) < 0$, and hence, by the continuity of g_i, there exists $\epsilon_2 > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for any $t \in (0, \epsilon_2)$ and $i \notin I(\mathbf{x}^*)$.
- Consequently,

$$
\begin{aligned}
& f(\mathbf{x}^* + t\mathbf{d}) < f(\mathbf{x}^*) \\
& g_i(\mathbf{x}^* + t\mathbf{d}) < 0, \quad i = 1, 2, \cdots, m
\end{aligned}
$$

  for all $t \in (0, \min\{\epsilon_1, \epsilon_2\})$
- Contradiction to the local optimality of $\mathbf{x}^*$.

$\square$

Theorem

Let $\mathbf{x}^*$ be a local minimum of the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \leq 0, \ i = 1, 2, \cdots, m \end{aligned}$$

where $f, g_1, \cdots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$. Then there exist multipliers $\lambda_0, \lambda_1, \cdots, \lambda_m \geq 0$, which are not all zeros, such that

$$\lambda_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \ i = 1, 2, \cdots, m$$

Proof.

- The following system is infeasible

$$(S) \quad \nabla f(\mathbf{x}^*)^\top \mathbf{d} < 0,\ \nabla g_i(\mathbf{x}^*)^\top \mathbf{d} < 0,\ i \in I(\mathbf{x}^*) = \{i_1, \ldots, i_k\}$$

- System (S) is the same as $\mathbf{Ad} < 0$ where $\mathbf{A} = (\nabla f(\mathbf{x}^*), \nabla g_{i_1}(\mathbf{x}^*), \cdots \nabla g_{i_k}(\mathbf{x}^*))^\top$
- By Gordan's theorem of alternative, system (S) is infeasible if and only if there exists a vector $\eta = (\lambda_0, \lambda_{i_1}, \cdots, \lambda_{i_k})^\top \neq \mathbf{0}$ such that $\mathbf{A}^\top \eta = \mathbf{0},\ \eta \geq \mathbf{0}$, which is the same as $\lambda_0 \nabla f(\mathbf{x}^*) + \sum_{i \in I(\mathbf{x}^*)} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.
- Define $\lambda_i = 0$ for any $i \notin I(\mathbf{x}^*)$, and we obtain that

$$\lambda_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0},\ \lambda_i g_i(\mathbf{x}^*) = 0,\ i = 1, 2, \cdots, m$$

$\square$

A major drawback of the Fritz-John conditions is, they allow λ_0 to be zero. Under an additional regularity condition, we can assume $\lambda_0 = 1$.

Theorem

Let $\mathbf{x}^*$ *be a local minimum of the problem*

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \leq 0, \ i = 1, 2, \cdots, m \end{aligned}$$

where $f, g_1, \cdots, g_m$ *are continuously differentiable functions over* $\mathbb{R}^n$ *. Suppose that the gradients of the active constraints* $\{\nabla g_i(\mathbf{x}^*)\}_{i \in I(\mathbf{x}^*)}$ *are linearly independent. Then there exist multipliers* $\lambda_1, \lambda_2, \cdots, \lambda_m \geq 0$, *such that*

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \ i = 1, 2, \cdots, m$$

Proof.

- By the Fritz-John conditions, there exists $\tilde{\lambda}_0, \tilde{\lambda}_1, \cdots, \tilde{\lambda}_m$, not all zeros, such that

$$\tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$
$$\tilde{\lambda}_i g_i(\mathbf{x}^*) = 0, \; i = 1, 2, \cdots, m$$

- $\tilde{\lambda}_0 \neq 0$ since otherwise, if $\tilde{\lambda}_0 = 0$, $\sum_{i \in I(\mathbf{x}^*)} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$, where not all the scalars $\tilde{\lambda}_i, i \in I(\mathbf{x}^*)$ are zeros, which is a contradiction to the regularity condition.
- $\tilde{\lambda}_0 > 0$. Defining $\lambda_i = \frac{\tilde{\lambda}_i}{\tilde{\lambda}_0}$, the result follows.

$\square$

Theorem

Let $\mathbf{x}^*$ *be a local minimum of the problem*

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & g_i(\mathbf{x}) \leq 0, \ i = 1, 2, \cdots, m \\ & h_j(\mathbf{x}) = 0, \ j = 1, 2, \cdots, p \end{array} \tag{1}$$

where $f, g_1, \cdots, g_m, h_1, h_2, \cdots, h_p$ *are continuously differentiable functions over* $\mathbb{R}^n$ *. Suppose that the gradients of the active constraints and the equality constraints:* $\{\nabla g_i(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*), i \in I(\mathbf{x}^*), j = 1, 2, \cdots, p\}$ *are linearly independent. Then there exist multipliers* $\lambda_1, \cdots, \lambda_m \geq 0, \mu_1, \mu_2, \cdots, \mu_p \in \mathbb{R}$, *such that*

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \ i = 1, 2, \cdots, m$$

Definition (KKT point)

Consider problem (1) where $g_1, \cdots, g_m, h_1, h_2, \cdots, h_p$ are continuously differentiable functions over $\mathbb{R}^n$. A feasible point $\mathbf{x}^*$ is called a KKT point if there exist $\lambda_1, \cdots, \lambda_m \geq 0, \mu_1, \mu_2, \cdots, \mu_p \in \mathbb{R}$ such that

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \;\; i = 1, 2, \cdots, m$$

Definition (Regularity)

A feasible point $\mathbf{x}^*$ is called regular if the set $\{\nabla g_i(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*), i \in I(\mathbf{x}^*), j = 1, 2, \cdots, p\}$ is linearly independent.

- The KKT theorem states that a necessary local optimality condition of a regular point is that it is a KKT point.
- The additional requirement of regularity is not required in linearly constrained problems in which no such assumption is needed.

Examples

1.
$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 = 1 \end{aligned}$$

2.
$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.t.} \quad & (x_1^2 + x_2^2 - 1)^2 = 0 \end{aligned}$$

In class.

Outline

Sufficiency of KKT Conditions in the Convex Case

In the convex case the KKT conditions are always sufficient.

Theorem

Let $\mathbf{x}^*$ be a feasible solution of the problem

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & g_i(\mathbf{x}) \leq 0, \ i = 1, 2, \cdots, m \\ & h_j(\mathbf{x}) = 0, \ j = 1, 2, \cdots, p \end{array} \tag{2}$$

where $f, g_1, \cdots, g_m$ are continuously differentiable convex functions over $\mathbb{R}^n$ and $h_1, h_2, \cdots, h_p$ are affine functions. Suppose that there exist multipliers $\lambda_1, \cdots, \lambda_m \geq 0, \mu_1, \mu_2, \cdots, \mu_p \in \mathbb{R}$, such that

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \ i = 1, 2, \cdots, m$$

Then $\mathbf{x}^*$ is the optimal solution of (2).

Proof.

- Let $\mathbf{x}$ be a feasible solution of (2). We will show that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.
- The function $s(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x})$ is convex.
- Since $\nabla s(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$, it follows that $\mathbf{x}^*$ is a minimizer of s over $\mathbb{R}^n$, and in particular $s(\mathbf{x}^*) \leq s(\mathbf{x})$.
- Thus,

$$\begin{aligned} f(\mathbf{x}^*) &= f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}^*) \\ &= s(\mathbf{x}^*) \\ &\leq s(\mathbf{x}) \\ &= f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}) \leq f(\mathbf{x}) \end{aligned}$$

$\square$

If constraints are convex, regularity can be replaced by Slater's condition.

Theorem

Let $\mathbf{x}^*$ *be a local minimum of the problem*

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & g_i(\mathbf{x}) \leq 0, \ i = 1, 2, \cdots, m \end{array} \tag{3}$$

where $f, g_1, \cdots, g_m$ *are continuously differentiable functions over* $\mathbb{R}^n$ *. In addition,* $g_1, \cdots, g_m$ *are convex over* $\mathbb{R}^n$*. Suppose* $\exists \hat{\mathbf{x}} \in \mathbb{R}^n$ *such that*

$$g_i(\hat{\mathbf{x}}) < 0, \quad i = 1, 2, \cdots, m$$

Then there exist multipliers $\lambda_1, \cdots, \lambda_m \geq 0$*, such that*

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0} \tag{4}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0, \ i = 1, 2, \cdots, m \tag{5}$$

Convex Constraints - Necessity under Slater's Condition

Proof.

- Since $\mathbf{x}^*$ is an optimal solution of (3), the Fritz-John conditions are satisfied: there exist $\tilde{\lambda}_0, \cdots, \tilde{\lambda}_m \geq 0$ not all zeros, such that

$$\tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$
$$\tilde{\lambda}_i g_i(\mathbf{x}^*) = 0, \ \ i = 1, 2, \cdots, m \tag{6}$$

- We will prove that $\tilde{\lambda}_0 > 0$, and then conditions (4) and (5) will be satisfied with $\lambda_i = \frac{\tilde{\lambda}_i}{\tilde{\lambda}_0}, i = 1, 2, ..., m$.
- Assume in contradiction that $\tilde{\lambda}_0 = 0$. Then

$$\sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) = \mathbf{0} \tag{7}$$

Proof Contd.

- By the gradient inequality,

$$0 > g_i(\hat{\mathbf{x}}) \geq g_i(\mathbf{x}^*) + \nabla g_i(\mathbf{x}^*)^\top (\hat{\mathbf{x}} - \mathbf{x}^*), \;\; i = 1, 2, \cdots, m$$

- Multiplying the i-th equation by $\tilde{\lambda}_i$ and summing over $i = 1, 2, \cdots, m$ we obtain

$$0 > \sum_{i=1}^{m} \tilde{\lambda}_i g_i(\mathbf{x}^*) + \Big[\sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) \Big]^\top (\hat{\mathbf{x}} - \mathbf{x}^*) \tag{8}$$

- Plugging the identities (7) and (6) into (8) we obtain the impossible statement that $0 > 0$, thus establishing the result.

$\square$

Examples

1.

$$\begin{aligned} \min \quad & x_1^2 - x_2 \\ \text{s.t.} \quad & x_2 = 0 \end{aligned}$$

2.

$$\begin{aligned} \min \quad & x_1^2 - x_2 \\ \text{s.t.} \quad & x_2^2 \leq 0 \end{aligned}$$

The optimal solution is $(x_1, x_2) = (0, 0)$. Satisfies KKT conditions for problem 1, but not for problem 2.
In class.

Definition (Generalized Slater's Condition)

Consider the system

$$g_i(\mathbf{x}) \leq 0, \; i = 1, 2, \cdots, m$$
$$h_j(\mathbf{x}) \leq 0, \; j = 1, 2, \cdots, p$$
$$s_k(\mathbf{x}) = 0, \; k = 1, 2, \cdots, q$$

where $g_i, i = 1, 2, \cdots, m$ are convex functions over $\mathbb{R}^n$ and $h_j, s_k, j = 1, 2, \cdots, p, k = 1, 2, \cdots, q$ are affine functions over $\mathbb{R}^n$. Then we say that the generalized Slater's condition is satisfied if there exists $\hat{\mathbf{x}} \in \mathbb{R}^n$ for which

$$g_i(\hat{\mathbf{x}}) < 0, \; i = 1, 2, \cdots, m$$
$$h_j(\hat{\mathbf{x}}) \leq 0, \; j = 1, 2, \cdots, p$$
$$s_k(\hat{\mathbf{x}}) = 0, \; k = 1, 2, \cdots, q$$

Necessity of KKT under Generalized Slater's Condition

Theorem

Let $\mathbf{x}^*$ *be an optimal solution of the problem*

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & g_i(\mathbf{x}) \leq 0,\ i = 1, 2, \cdots, m \\ & h_j(\mathbf{x}) \leq 0,\ j = 1, 2, \cdots, p \\ & s_k(\mathbf{x}) = 0,\ k = 1, 2, \cdots, q \end{array} \tag{9}$$

where $f, g_1, \cdots, g_m$ *are continuously differentiable functions and* $\{g_i\}$ *are convex functions over* $\mathbb{R}^n$, *and* $\{h_j\}, \{s_k\}$ *are affine. Suppose the generalized Slater's condition holds. Then there exist multipliers* $\lambda_1, \lambda_2 \cdots, \lambda_m, \eta_1, \eta_2, \cdots, \eta_p \geq 0, \mu_1, \mu_2, \cdots, \mu_q \in \mathbb{R}$ *such that*

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \eta_j \nabla h_j(\mathbf{x}^*) + \sum_{k=1}^{q} \mu_k \nabla s_k(\mathbf{x}^*) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0,\ i = 1, 2, \cdots, m$$

$$\eta_j h_j(\mathbf{x}^*) = 0,\ j = 1, 2, \cdots, p$$

Example

$$
\begin{aligned}
\min \quad & 4x_1^2 + x_2^2 - x_1 - 2x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 \leq 1 \\
& x_1^2 \leq 1
\end{aligned}
$$

In class.

Outline

$$(CLS) \quad \begin{aligned} \min \quad & \|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{s.t.} \quad & \|\mathbf{x}\|^2 \leq \alpha \end{aligned}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ has full column rank, $\mathbf{b} \in \mathbb{R}^m, \alpha > 0$.

- Problem (CLS) is a convex problem and satisfies Slater's condition.
- Lagrangian: $L(\mathbf{x}, \lambda) = \|\mathbf{Ax} - \mathbf{b}\|^2 + \lambda(\|\mathbf{x}\|^2 - \alpha) \quad (\lambda \geq 0)$
- KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{x}} L = 2\mathbf{A}^\top(\mathbf{Ax} - \mathbf{b}) + 2\lambda\mathbf{x} = 0 \\ \lambda(\|\mathbf{x}\|^2 - \alpha) = 0 \\ \|\mathbf{x}\|^2 \leq \alpha,\ \lambda \geq 0 \end{aligned}$$

- If $\lambda = 0$, then by the first equation

$$\mathbf{x} = \mathbf{x}_{LS} \equiv (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{b}.$$

Optimal iff $\|\mathbf{x}_{LS}\|^2 \leq \alpha$.

- On the other hand, if $\|\mathbf{x}_{LS}\|^2 > \alpha$, then necessarily $\lambda > 0$. By the C-S condition we have that $\|\mathbf{x}\|^2 = \alpha$ and the first equation implies that

$$\mathbf{x} = \mathbf{x}_\lambda \equiv (\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{b}$$

The multiplier $\lambda > 0$ should be chosen to satisfy $\|\mathbf{x}_\lambda\|^2 = \alpha$, that is, λ is the solution of

$$f(\lambda) = \|(\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{b}\|^2 - \alpha = 0$$

- $f(0) = \|(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}\|^2 - \alpha = \|\mathbf{x}_{LS}\|^2 - \alpha > 0$, f strictly decreasing and $f(\lambda) \to -\alpha$ as $\lambda \to \infty$.
- Conclusion: the optimal solution of the CLS problem is given by

$$\mathbf{x} = \begin{cases} \mathbf{x}_{LS} & \|\mathbf{x}_{LS}\|^2 \le \alpha \\ (\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{b} & \|\mathbf{x}_{LS}\|^2 > \alpha \end{cases}$$

where λ is the unique root of $f(\lambda)$ over $(0, \infty)$.

Outline

Theorem

Consider the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \le 0, \ i = 1, 2, \cdots, m \end{aligned}$$

where $f, g_1, \cdots, g_m$ are twice continuously differentiable functions over $\mathbb{R}^n$. Let $\mathbf{x}^*$ be a local minimum, and suppose that $\mathbf{x}^*$ is regular meaning that $\{\nabla g_i(\mathbf{x}^*)\}_{i \in I(\mathbf{x}^*)}$ are linearly independent. Then $\exists\ \lambda_1, \lambda_2 \cdots, \lambda_m \ge 0$ such that

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda) &= \mathbf{0} \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \ i = 1, 2, \cdots, m \end{aligned}$$

and $\mathbf{d}^\top \nabla^2_{\mathbf{xx}} L(\mathbf{x}^*, \lambda)\mathbf{d} \ge 0$ for all $\mathbf{d} \in \Lambda(\mathbf{x}^*)$ where

$$\Lambda(\mathbf{x}^*) \equiv \{\mathbf{d} \in \mathbb{R}^n : \nabla g_i(\mathbf{x}^*)^\top \mathbf{d} = 0, \ i \in I(\mathbf{x}^*)\}.$$

See the proof of Theorem 11.18 in the textbook.

Theorem

Consider the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \leq 0,\ i = 1, 2, \cdots, m \\ & h_j(\mathbf{x}) = 0,\ j = 1, 2, \cdots, p \end{aligned}$$

where $f, g_1, \cdots, g_m, h_1, \cdots, h_p$ *are twice continuously differentiable functions. Let* $\mathbf{x}^*$ *be a local minimum and suppose that* $\mathbf{x}^*$ *is regular meaning that the set* $\{\nabla g_i(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*), i \in I(\mathbf{x}^*), j = 1, 2, \cdots, p\}$ *is linearly independent. Then* $\exists\ \lambda_1, \lambda_2, \cdots, \lambda_m \geq 0$ *and* $\mu_1, \mu_2, \cdots, \mu_p \in \mathbb{R}$ *such that*

$$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda, \mu) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{x}^*) = 0,\ i = 1, 2, \cdots, m$$

and $\mathbf{d}^\top \nabla^2_{\mathbf{xx}} L(\mathbf{x}^*, \lambda, \mu)\mathbf{d} \geq 0$ *for all* $\mathbf{d} \in \Lambda(\mathbf{x}^*) \equiv \{\mathbf{d} \in \mathbb{R}^n : \nabla g_i(\mathbf{x}^*)^\top \mathbf{d} = 0,\ \nabla h_j(\mathbf{x}^*)^\top \mathbf{d} = 0,\ i \in I(\mathbf{x}^*),\ j = 1, 2, \cdots, p\}$.

Outline

The Trust Region Subproblem (TRS) is the problem consisting of minimizing an indefinite quadratic function subject to an l_2-norm constraint:

$$(TRS): \quad \min\{f(\mathbf{x}) \equiv \mathbf{x}^\top \mathbf{A}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c : \ \|\mathbf{x}\|^2 \leq \alpha\}$$

where $\mathbf{A} = \mathbf{A}^\top \in \mathbb{R}^{n\times n}, \mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Although the problem is nonconvex, it possesses necessary and sufficient optimality conditions.

Theorem

A vector $\mathbf{x}^*$ is an optimal solution of problem (TRS) if and only if there exists $\lambda^* \geq 0$ such that

$$(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x}^* = -\mathbf{b} \tag{10}$$

$$\|\mathbf{x}^*\|^2 \leq \alpha \tag{11}$$

$$\lambda^*(\|\mathbf{x}^*\|^2 - \alpha) = 0 \tag{12}$$

$$\mathbf{A} + \lambda^* \mathbf{I} \succeq \mathbf{0} \tag{13}$$

Proof.

Sufficiency:

- Assume that $\mathbf{x}^*$ satisfies (10)-(13) for some $\lambda^* \geq 0$.
- Define the function

$$\begin{aligned} h(\mathbf{x}) &= \mathbf{x}^\top \mathbf{A}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c + \lambda^*(\|\mathbf{x}\|^2 - \alpha) \\ &= \mathbf{x}^\top (\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c - \alpha\lambda^* \end{aligned} \tag{14}$$

- Then by (13) we have that h is a convex quadratic function. By (10) it follows that $\nabla h(\mathbf{x}^*) = 0$, which implies that $\mathbf{x}^*$ is the unconstrained minimizer of h over $\mathbb{R}^n$.
- Let $\mathbf{x}$ be a feasible point, i.e., $\|\mathbf{x}\|^2 \leq \alpha$. Then

$$\begin{aligned} f(\mathbf{x}) &\geq f(\mathbf{x}) + \lambda^*(\|\mathbf{x}\|^2 - \alpha) && (\lambda^* \geq 0,\ \|\mathbf{x}\|^2 - \alpha \leq 0) \\ &= h(\mathbf{x}) && (\text{by } (14)) \\ &\geq h(\mathbf{x}^*) && (\mathbf{x}^* \text{ is the minimizer of } h) \\ &= f(\mathbf{x}^*) + \lambda^*(\|\mathbf{x}^*\|^2 - \alpha) \\ &= f(\mathbf{x}^*) && (\text{by } (12)) \end{aligned}$$

Proof Contd.

Necessity:

- Regularity always satisfied. Suppose $\mathbf{x}^*$ is a minimizer of (TRS). If $\|\mathbf{x}^*\|^2 < \alpha$, done. If $\|\mathbf{x}^*\|^2 = \alpha$, by SONC, $\exists \lambda^* \geq 0$ such that

$$(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x}^* = -\mathbf{b} \tag{15}$$

$$\|\mathbf{x}^*\|^2 \leq \alpha \tag{16}$$

$$\lambda^*(\|\mathbf{x}^*\|^2 - \alpha) = 0 \tag{17}$$

$$\mathbf{d}^\top(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{d} \geq 0 \quad \text{for all } \mathbf{d} \text{ satisfying } \mathbf{d}^\top \mathbf{x}^* = 0 \tag{18}$$

- Need to show that (18) is true for any $\mathbf{d}$.
- Suppose on the contrary: $\exists \mathbf{d}$ s.t. $\mathbf{d}^\top \mathbf{x}^* > 0$ and $\mathbf{d}^\top(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{d} < 0$.
- Consider $\bar{\mathbf{x}} = \mathbf{x}^* + t\mathbf{d}$, where $t = -2\frac{\mathbf{d}^\top \mathbf{x}^*}{\|\mathbf{d}\|^2}$. $\bar{\mathbf{x}}$ is feasible since

$$\begin{aligned} \|\bar{\mathbf{x}}\|^2 &= \|\mathbf{x}^* + t\mathbf{d}\|^2 = \|\mathbf{x}^*\|^2 + 2t\mathbf{d}^\top \mathbf{x}^* + t^2\|\mathbf{d}\|^2 \\ &= \|\mathbf{x}^*\|^2 - 4\frac{(\mathbf{d}^\top \mathbf{x}^*)^2}{\|\mathbf{d}\|^2} + 4\frac{(\mathbf{d}^\top \mathbf{x}^*)^2}{\|\mathbf{d}\|^2} = \|\mathbf{x}^*\|^2 \leq \alpha. \end{aligned}$$

Proof Contd.

- In addition,

$$
\begin{aligned}
f(\bar{\mathbf{x}}) &= \bar{\mathbf{x}}^\top \mathbf{A}\bar{\mathbf{x}} + 2\mathbf{b}^\top \bar{\mathbf{x}} + c \\
&= (\mathbf{x}^* + t\mathbf{d})^\top \mathbf{A}(\mathbf{x}^* + t\mathbf{d}) + 2\mathbf{b}^\top(\mathbf{x}^* + t\mathbf{d}) + c \\
&= \underbrace{(\mathbf{x}^*)^\top \mathbf{A}\mathbf{x}^* + 2\mathbf{b}^\top \mathbf{x}^* + c}_{f(\mathbf{x}^*)} + t^2\mathbf{d}^\top \mathbf{A}\mathbf{d} + 2t\mathbf{d}^\top(\mathbf{A}\mathbf{x}^* + \mathbf{b}) \\
&= f(\mathbf{x}^*) + t^2\mathbf{d}^\top(\mathbf{A} + \lambda^*\mathbf{I})\mathbf{d} + 2t\mathbf{d}^\top \underbrace{((\mathbf{A} + \lambda^*\mathbf{I})\mathbf{x}^* + \mathbf{b})}_{=0 \text{ by } (15)} \\
&\quad - \lambda^* t \underbrace{[t\|\mathbf{d}\|^2 + 2\mathbf{d}^\top \mathbf{x}^*]}_{=0} \\
&= f(\mathbf{x}^*) + t^2\mathbf{d}^\top(\mathbf{A} + \lambda^*\mathbf{I})\mathbf{d} \\
&< f(\mathbf{x}^*)
\end{aligned}
$$

which is a contradiction to the optimality of $\mathbf{x}^*$. □

Outline

Consider the approximate set of linear equations:

$$\mathbf{Ax} \approx \mathbf{b}$$

- In the Least Squares (LS) approach we only assume $\mathbf{b}$ is subjected to noise.

$$\begin{aligned} \min_{\mathbf{w},\mathbf{x}} \quad & \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} + \mathbf{w} \\ & \mathbf{w} \in \mathbb{R}^m \end{aligned}$$

- In the Total Least Squares (TLS) we assume that both the RHS vector $\mathbf{b}$ and the model matrix $\mathbf{A}$ are subjected to noise

$$(TLS) \quad \begin{aligned} \min_{\mathbf{E},\mathbf{w},\mathbf{x}} \quad & \|\mathbf{E}\|_F^2 + \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & (\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w} \\ & \mathbf{E} \in \mathbb{R}^{m \times n}, \ \mathbf{w} \in \mathbb{R}^m \end{aligned}$$

The TLS problem – as formulated – seems like a difficult nonconvex problem. Will see that it can be solved efficiently.

- Fixing $\mathbf{x}$, we will solve the problem

$$(P_{\mathbf{x}}) \quad \begin{aligned} \min_{\mathbf{E},\mathbf{w}} \quad & \|\mathbf{E}\|_F^2 + \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & (\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w} \end{aligned}$$

- The KKT conditions are necessary and sufficient for problem $(P_{\mathbf{x}})$.
- Lagrangian: $L(\mathbf{E}, \mathbf{w}, \lambda) = \|\mathbf{E}\|_F^2 + \|\mathbf{w}\|^2 + 2\lambda^\top[(\mathbf{A}+\mathbf{E})\mathbf{x} - \mathbf{b} - \mathbf{w}]$
- By the KKT conditions, $(\mathbf{E}, \mathbf{w})$ is an optimal solution of $(P_{\mathbf{x}})$ if and only if there exists $\lambda \in \mathbb{R}^m$ such that

$$2\mathbf{E} + 2\lambda\mathbf{x}^\top = \mathbf{0} \qquad (\nabla_{\mathbf{E}} L = \mathbf{0}) \qquad (19)$$

$$2\mathbf{w} - 2\lambda = \mathbf{0} \qquad (\nabla_{\mathbf{w}} L = \mathbf{0}) \qquad (20)$$

$$(\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w} \qquad (\text{feasibility}) \qquad (21)$$

- By (19), (20) and (21), $\mathbf{E} = -\lambda\mathbf{x}^\top, \mathbf{w} = \lambda$ and $\lambda = \frac{\mathbf{A}\mathbf{x}-\mathbf{b}}{\|\mathbf{x}\|^2+1}$. Plugging this into the objectve function, a reduced formulation in the variables $\mathbf{x}$ is obtained.

$$(TLS') \quad \min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{Ax}-\mathbf{b}\|^2}{\|\mathbf{x}\|^2+1}$$

Theorem

$\mathbf{x}$ is an optimal solution of (TLS') if and only if $(\mathbf{x},\mathbf{E},\mathbf{w})$ is an optimal solution of (TLS) where $\mathbf{E}=-\frac{(\mathbf{Ax}-\mathbf{b})\mathbf{x}^\top}{\|\mathbf{x}\|^2+1}$ and $\mathbf{w}=\frac{\mathbf{Ax}-\mathbf{b}}{\|\mathbf{x}\|^2+1}$

- Still a nonconvex problem.
- Resembles the problem of minimizing the Rayleigh quotient.

Under a rather mild condition, the optimal solution of (TLS') can be derived via a homogenization argument.

- (TLS') is the same as

$$\min_{\mathbf{x}\in\mathbb{R}^n, t\in\mathbb{R}} \left\{ \frac{\|\mathbf{A}\mathbf{x} - t\mathbf{b}\|^2}{\|\mathbf{x}\|^2 + t^2} : t = 1 \right\}$$

- The same as (denoting $\mathbf{y} = \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}$):

$$f^* = \min_{\mathbf{y}\in\mathbb{R}^{n+1}} \left\{ \frac{\mathbf{y}^\top \mathbf{B}\mathbf{y}}{\|\mathbf{y}\|^2} : y_{n+1} = 1 \right\} \tag{22}$$

where

$$\mathbf{B} = \begin{pmatrix} \mathbf{A}^\top \mathbf{A} & -\mathbf{A}^\top \mathbf{b} \\ -\mathbf{b}^\top \mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix}$$

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

$$g^* = \min_{\mathbf{y} \in \mathbb{R}^{n+1}} \left\{ \frac{\mathbf{y}^\top \mathbf{B} \mathbf{y}}{\|\mathbf{y}\|^2} : \mathbf{y} \neq \mathbf{0} \right\} \tag{23}$$

Lemma

Let $\mathbf{y}^*$ be an optimal solution of (23) and assume that $y^*_{n+1} \neq 0$. Then $\tilde{\mathbf{y}} = \frac{1}{y^*_{n+1}} \mathbf{y}^*$ is an optimal solution of (22).

Proof.

- $f^* \geq g^*$.
- $\tilde{\mathbf{y}}$ is feasible for (22) and we have

$$f^* \leq \frac{\tilde{\mathbf{y}}^\top \mathbf{B} \tilde{\mathbf{y}}}{\|\tilde{\mathbf{y}}\|^2} = \frac{\frac{1}{(y^*_{n+1})^2} (\mathbf{y}^*)^\top \mathbf{B} \mathbf{y}^*}{\frac{1}{(y^*_{n+1})^2} \|\mathbf{y}^*\|^2} = \frac{(\mathbf{y}^*)^\top \mathbf{B} \mathbf{y}^*}{\|\mathbf{y}^*\|^2} = g^*$$

- Therefore, $\tilde{\mathbf{y}}$ is an optimal solution of both (22) and (23).

$\square$

Main Result on TLS

Theorem

Assume that the following condition holds:

$$\lambda_{\min}(\mathbf{B}) < \lambda_{\min}(\mathbf{A}^\top \mathbf{A}) \tag{24}$$

where

$$\mathbf{B} = \begin{pmatrix} \mathbf{A}^\top \mathbf{A} & -\mathbf{A}^\top \mathbf{b} \\ -\mathbf{b}^\top \mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix}$$

Then the optimal solution of problem (TLS') is given by $\frac{1}{y_{n+1}^*}\mathbf{v}$*, where* $\mathbf{y}^* = \begin{pmatrix} \mathbf{v} \\ y_{n+1}^* \end{pmatrix}$ *is an eigenvector corresponding to the minimum eigenvalue of* $\mathbf{B}$*.*

Proof.

- All we need to prove is that under condition (24), an optimal solution $\mathbf{y}^*$ of (23) must satisfy $y^*_{n+1} \neq 0$.
- Assume on the contrary that $y^*_{n+1} = 0$. Then

$$\lambda_{\min}(\mathbf{B}) = \frac{(\mathbf{y}^*)^\top \mathbf{B}\mathbf{y}^*}{\|\mathbf{y}^*\|^2} = \frac{\mathbf{v}^\top \mathbf{A}^\top \mathbf{A}\mathbf{v}}{\|\mathbf{v}\|^2} \geq \lambda_{\min}(\mathbf{A}^\top \mathbf{A})$$

which is a contradiction to (24).

□