Lecture 11 - The Karush-Kuhn-Tucker Conditions

- 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.

Lecture 11 - The Karush-Kuhn-Tucker Conditions

- 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

$$\begin{array}{ll} \min & h(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C, \end{array}$$

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

The Basic Necessary Condition - No Feasible Descent Directions

Lemma. Consider the problem

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

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

The Basic Necessary Condition - No Feasible Descent Directions

Lemma. Consider the problem

$$(\mathrm{G}) \quad \begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C, \end{array}$$

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

Proof.

- By contradiction, assume that there exists a vector $\mathbf{d}$ and $\varepsilon_1 > 0$ such that $\mathbf{x} + t\mathbf{d} \in C$ for all $t \in [0, \varepsilon_1]$ and $\nabla f(\mathbf{x}^*)^T\mathbf{d} < 0$.

The Basic Necessary Condition - No Feasible Descent Directions

Lemma. Consider the problem

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

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

Proof.

- By contradiction, assume that there exists a vector $\mathbf{d}$ and $\varepsilon_1 > 0$ such that $\mathbf{x} + t\mathbf{d} \in C$ for all $t \in [0, \varepsilon_1]$ and $\nabla f(\mathbf{x}^*)^T \mathbf{d} < 0$.
- By definition of the directional derivative there exists $\varepsilon_2 < \varepsilon_1$ such that $f(\mathbf{x}^* + t\mathbf{d}) < f(\mathbf{x}^*)$ for all $t \in [0, \varepsilon_2] \Rightarrow$ contradiction to the local optimality of $\mathbf{x}^*$.

Consequence

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

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

where $f, g_1, \ldots, g_m$ are continuously differentiable functions 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

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

Proof

- Suppose that $\mathbf{d}$ satisfies the system of inequalities.

Proof

- Suppose that **d** satisfies the system of inequalities.
- Then $\exists \varepsilon_1 > 0$ 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 $t \in (0, \varepsilon_1)$ and $i \in I(\mathbf{x}^*)$.

Proof

- Suppose that $\mathbf{d}$ satisfies the system of inequalities.
- Then $\exists \varepsilon_1 > 0$ 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 $t \in (0, \varepsilon_1)$ 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 $\varepsilon_2 > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for any $t \in (0, \varepsilon_2)$ and $i \notin I(\mathbf{x}^*)$.

Proof

- Suppose that $\mathbf{d}$ satisfies the system of inequalities.
- Then $\exists \varepsilon_1 > 0$ 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 $t \in (0, \varepsilon_1)$ 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 $\varepsilon_2 > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for any $t \in (0, \varepsilon_2)$ and $i \notin I(\mathbf{x}^*)$.
- Consequently,

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

for all $t \in (0, \min\{\varepsilon_1, \varepsilon_2\})$.

Proof

- Suppose that $\mathbf{d}$ satisfies the system of inequalities.
- Then $\exists \varepsilon_1 > 0$ 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 $t \in (0, \varepsilon_1)$ 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 $\varepsilon_2 > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for any $t \in (0, \varepsilon_2)$ and $i \notin I(\mathbf{x}^*)$.
- Consequently,

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

  for all $t \in (0, \min\{\varepsilon_1, \varepsilon_2\})$.
- A contradiction to the local optimality of $\mathbf{x}^*$.

The Fritz-John Necessary Condition

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

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

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

$$\begin{aligned} \lambda_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Proof of Fritz-John Conditions

- The following system is infeasible

$$\text{(S)} \quad \nabla f(\mathbf{x}^*)^T \mathbf{d} < 0, \nabla g_i(\mathbf{x}^*)^T \mathbf{d} < 0, i \in I(\mathbf{x}^*)$$

Proof of Fritz-John Conditions

- The following system is infeasible

$$\text{(S)} \quad \nabla f(\mathbf{x}^*)^T \mathbf{d} < 0, \nabla g_i(\mathbf{x}^*)^T \mathbf{d} < 0, i \in I(\mathbf{x}^*)$$

- System (S) is the same as $\mathbf{Ad} < \mathbf{0}$ where $\mathbf{A} = \begin{pmatrix} \nabla f(\mathbf{x}^*)^T \\ \nabla g_{i_1}(\mathbf{x}^*)^T \\ \vdots \\ \nabla g_{i_k}(\mathbf{x}^*)^T \end{pmatrix}$

Proof of Fritz-John Conditions

- The following system is infeasible

$$\text{(S)} \quad \nabla f(\mathbf{x}^*)^T \mathbf{d} < 0, \nabla g_i(\mathbf{x}^*)^T \mathbf{d} < 0, i \in I(\mathbf{x}^*)$$

- System (S) is the same as $\mathbf{Ad} < \mathbf{0}$ where $\mathbf{A} = \begin{pmatrix} \nabla f(\mathbf{x}^*)^T \\ \nabla g_{i_1}(\mathbf{x}^*)^T \\ \vdots \\ \nabla g_{i_k}(\mathbf{x}^*)^T \end{pmatrix}$
- By Gordan's theorem of alternative, system (S) is infeasible if and only if there exists a vector $\boldsymbol{\eta} = (\lambda_0, \lambda_{i_1}, \ldots, \lambda_{i_k})^T \neq \mathbf{0}$ such that

$$\mathbf{A}^T \boldsymbol{\eta} = \mathbf{0}, \boldsymbol{\eta} \geq \mathbf{0},$$

Proof of Fritz-John Conditions

- The following system is infeasible

$$\text{(S)} \quad \nabla f(\mathbf{x}^*)^T \mathbf{d} < 0, \nabla g_i(\mathbf{x}^*)^T \mathbf{d} < 0, i \in I(\mathbf{x}^*)$$

- System (S) is the same as $\mathbf{Ad} < \mathbf{0}$ where $\mathbf{A} = \begin{pmatrix} \nabla f(\mathbf{x}^*)^T \\ \nabla g_{i_1}(\mathbf{x}^*)^T \\ \vdots \\ \nabla g_{i_k}(\mathbf{x}^*)^T \end{pmatrix}$
- By Gordan's theorem of alternative, system (S) is infeasible if and only if there exists a vector $\boldsymbol{\eta} = (\lambda_0, \lambda_{i_1}, \ldots, \lambda_{i_k})^T \neq \mathbf{0}$ such that

$$\mathbf{A}^T \boldsymbol{\eta} = \mathbf{0}, \boldsymbol{\eta} \geq \mathbf{0},$$

- which is the same as $\lambda_0 \nabla f(\mathbf{x}^*) + \sum_{i \in I(\mathbf{x}^*)} \lambda_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.

Proof of Fritz-John Conditions

- The following system is infeasible

$$\text{(S)} \quad \nabla f(\mathbf{x}^*)^T \mathbf{d} < 0, \nabla g_i(\mathbf{x}^*)^T \mathbf{d} < 0, i \in I(\mathbf{x}^*)$$

- System (S) is the same as $\mathbf{Ad} < \mathbf{0}$ where $\mathbf{A} = \begin{pmatrix} \nabla f(\mathbf{x}^*)^T \\ \nabla g_{i_1}(\mathbf{x}^*)^T \\ \vdots \\ \nabla g_{i_k}(\mathbf{x}^*)^T \end{pmatrix}$
- By Gordan's theorem of alternative, system (S) is infeasible if and only if there exists a vector $\boldsymbol{\eta} = (\lambda_0, \lambda_{i_1}, \ldots, \lambda_{i_k})^T \neq \mathbf{0}$ such that

$$\mathbf{A}^T \boldsymbol{\eta} = \mathbf{0}, \boldsymbol{\eta} \geq \mathbf{0},$$

- 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

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

The KKT Conditions for Inequality Constrained Problems

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

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

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

where $f, g_1, \ldots, g_m$ 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 $\lambda_1, \lambda_2 \ldots, \lambda_m \geq 0$ such that

$$\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Proof of the KKT Conditions for Inequality Constrained Problems

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

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Proof of the KKT Conditions for Inequality Constrained Problems

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

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

- $\tilde{\lambda}_0 \neq 0$ 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.

Proof of the KKT Conditions for Inequality Constrained Problems

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

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

- $\tilde{\lambda}_0 \neq 0$ 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.

KKT Conditions for Inequality/Equality Constrained Problems

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

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

where $f, g_1, \ldots, g_m, h_1, h_2, \ldots, h_p$ are continuously differentiable functions over $\mathbb{R}^n$. Suppose that the gradients of the active constraints and the equality constraints: $\{\nabla g_i(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*), i \in I(\mathbf{x}^*), j = 1, 2, \ldots, p\}$ are linearly independent. Then there exist multipliers $\lambda_1, \lambda_2 \ldots, \lambda_m \geq 0, \mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Terminology

Definition (KKT point) Consider problem (1) where $f, g_1, \ldots, g_m, h_1, h_2, \ldots, h_p$ are continuously differentiable functions over $\mathbb{R}^n$. A feasible point $\mathbf{x}^*$ is called a KKT point if there exist $\lambda_1, \lambda_2 \ldots, \lambda_m \geq 0, \mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

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

Terminology

Definition (KKT point) Consider problem (1) where $f, g_1, \ldots, g_m, h_1, h_2, \ldots, h_p$ are continuously differentiable functions over $\mathbb{R}^n$. A feasible point $\mathbf{x}^*$ is called a KKT point if there exist $\lambda_1, \lambda_2 \ldots, \lambda_m \geq 0, \mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Definition (regularity) A feasible point $\mathbf{x}^*$ is called regular if the set $\{\nabla g_i(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*), i \in I(\mathbf{x}^*), j = 1, 2, \ldots, p\}$ 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. $$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.t.} & x_1^2 + x_2^2 = 1. \end{array}$$

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

In class

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

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

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

$$
\begin{aligned}
\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0}, \\
\lambda_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m.
\end{aligned}
$$

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

Proof

- Let $\mathbf{x}$ be a feasible solution of (2). We will show that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.

Proof

- Let $\mathbf{x}$ be a feasible solution of (2). We will show that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.
- The function $s(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^m \mu_i h_i(\mathbf{x})$ is convex.

Proof

- Let $\mathbf{x}$ be a feasible solution of (2). We will show that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.
- The function $s(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^m \mu_i h_i(\mathbf{x})$ 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 $s(\mathbf{x}^*) \leq s(\mathbf{x})$.

Proof

- Let $\mathbf{x}$ be a feasible solution of (2). We will show that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.
- The function $s(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^m \mu_i h_i(\mathbf{x})$ 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,

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

Convex Constraints - Necessity under Slater's Condition

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

Theorem (necessity of the KKT conditions under Slater's condition) Let $\mathbf{x}^*$ be a local optimal solution of the problem

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

where $f, g_1, \ldots, g_m$ are continuously differentiable over $\mathbb{R}^n$. In addition, $g_1, g_2, \ldots, g_m$ are convex over $\mathbb{R}^n$. Suppose $\exists \hat{\mathbf{x}} \in \mathbb{R}^n$ such that

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

Then there exist multipliers $\lambda_1, \lambda_2 \ldots, \lambda_m \geq 0$ such that

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

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

Proof

- Since $\mathbf{x}^*$ is an optimal solution of (3), the Fritz-John conditions are satisfied: there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$ not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{6}$$

Proof

- Since $\mathbf{x}^*$ is an optimal solution of (3), the Fritz-John conditions are satisfied: there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$ not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{6}$$

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

Proof

- Since $\mathbf{x}^*$ is an optimal solution of (3), the Fritz-John conditions are satisfied: there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$ not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{6}$$

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

- Since $\mathbf{x}^*$ is an optimal solution of (3), the Fritz-John conditions are satisfied: there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$ not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\mathbf{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\mathbf{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{6}$$

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

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

- By the gradient inequality,

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

Proof Contd.

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

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

Proof Contd.

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

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

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

Examples

1.
$$\begin{array}{ll} \min & x_1^2 - x_2 \\ \text{s.t.} & x_2 = 0. \end{array}$$

.

2.
$$\begin{array}{ll} \min & x_1^2 - x_2 \\ \text{s.t.} & x_2^2 \le 0. \end{array}$$

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

The Convex Case - Generalized Slater's Condition

Definition (Generalized Slater's Condition) Consider the system

$$\begin{aligned} g_i(\mathbf{x}) &\le 0, \quad i = 1, 2, \ldots, m, \\ h_j(\mathbf{x}) &\le 0, \quad j = 1, 2, \ldots, p, \\ s_k(\mathbf{x}) &= 0, \quad k = 1, 2, \ldots, q, \end{aligned}$$

where $g_i, i = 1, 2, \ldots, m$ are convex functions over $\mathbb{R}^n$ and $h_j, s_k, j = 1, 2, \ldots, p, k = 1, 2, \ldots, q$ 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

$$\begin{aligned} g_i(\hat{\mathbf{x}}) &< 0, \quad i = 1, 2, \ldots, m, \\ h_j(\hat{\mathbf{x}}) &\le 0, \quad j = 1, 2, \ldots, p, \\ s_k(\hat{\mathbf{x}}) &= 0, \quad k = 1, 2, \ldots, q, \end{aligned}$$

Necessity of KKT under Generalized Slater

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

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

where $f, g_1, \ldots, g_m$ are continuously differentiable convex functions and $h_j, s_k, j = 1, 2, \ldots, p, k = 1, 2, \ldots, q$ are affine. Suppose that the generalized Slater's condition is satisfied. Then there exist multipliers $\lambda_1, \lambda_2 \ldots, \lambda_m, \eta_1, \eta_2, \ldots, \eta_p \geq 0, \mu_1, \mu_2, \ldots, \mu_q \in \mathbb{R}$ such that

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

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

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

Example

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

In class

Constrained Least Squares

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

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.

Constrained Least Squares

$$\text{(CLS)} \quad \begin{array}{ll} \min & \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2, \\ \text{s.t.} & \|\mathbf{x}\|^2 \le \alpha, \end{array}$$

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: $L(\mathbf{x}, \lambda) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \lambda(\|\mathbf{x}\|^2 - \alpha). \quad (\lambda \ge 0)$

Constrained Least Squares

$$\text{(CLS)} \quad \begin{array}{ll} \min & \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2, \\ \text{s.t.} & \|\mathbf{x}\|^2 \le \alpha, \end{array}$$

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: $L(\mathbf{x}, \lambda) = \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \lambda(\|\mathbf{x}\|^2 - \alpha). \quad (\lambda \ge 0)$
- KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{x}} L = 2\mathbf{A}^T(\mathbf{A}\mathbf{x}-\mathbf{b}) + 2\lambda\mathbf{x} &= 0, \\ \lambda(\|\mathbf{x}\|^2 - \alpha) &= 0, \\ \|\mathbf{x}\|^2 &\le \alpha, \lambda \ge 0. \end{aligned}$$

Constrained Least Squares

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

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:

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

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

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

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

Constrained Least Squares Contd.

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

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

Constrained Least Squares Contd.

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

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

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

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

Constrained Least Squares Contd.

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

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

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

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

- $f(0) = \|(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}\|^2 - \alpha = \|\mathbf{x}_{\rm LS}\|^2 - \alpha > 0$, f strictly decreasing and $f(\lambda) \to -\alpha$ as $\lambda \to \infty$.

Constrained Least Squares Contd.

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

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

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

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

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

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

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

Second Order Necessary Optimality Conditions

Theorem. Consider the problem

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & f_i(\mathbf{x}) \le 0, \quad i = 1, 2, \ldots, m, \end{array}$$

where $f_0, f_1, \ldots, f_m$ are continuously differentiable over $\mathbb{R}^n$. Let $\mathbf{x}^*$ be a local minimum, and suppose that $\mathbf{x}^*$ is regular meaning that $\{\nabla f_i(\mathbf{x}^*)\}_{i \in I(\mathbf{x}^*)}$ are linearly independent. Then $\exists \lambda_1, \lambda_2, \ldots, \lambda_m \ge 0$ such that

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

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

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

See proof of Theorem 11.18 in the book

Second Order Necessary Optimality Conditions for Inequality/Equality Constrained Problems

Theorem. Consider the problem

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

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

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

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

Optimality Conditions for the Trust Region Subproblem

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

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

where $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{n\times n}, \mathbf{b} \in \mathbb{R}^n$ 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

$$\begin{aligned} (\mathbf{A} + \lambda^*\mathbf{I})\mathbf{x}^* &= -\mathbf{b} && (10)\\ \|\mathbf{x}^*\|^2 &\leq \alpha, && (11)\\ \lambda^*(\|\mathbf{x}^*\|^2 - \alpha) &= 0, && (12)\\ \mathbf{A} + \lambda^*\mathbf{I} &\succeq \mathbf{0}. && (13) \end{aligned}$$

Proof

Sufficiency:

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

Proof

Sufficiency:

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

$$h(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}+2\mathbf{b}^T\mathbf{x}+c+\lambda^*(\|\mathbf{x}\|^2-\alpha) = \mathbf{x}^T(\mathbf{A}+\lambda^*\mathbf{I})\mathbf{x}+2\mathbf{b}^T\mathbf{x}+c-\alpha\lambda^*. \quad (14)$$

Proof

Sufficiency:

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

$$h(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}+2\mathbf{b}^T\mathbf{x}+c+\lambda^*(\|\mathbf{x}\|^2-\alpha) = \mathbf{x}^T(\mathbf{A}+\lambda^*\mathbf{I})\mathbf{x}+2\mathbf{b}^T\mathbf{x}+c-\alpha\lambda^*. \quad (14)$$

- 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 $\mathbb{R}^n$.

Proof

Sufficiency:

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

$$h(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c + \lambda^*(\|\mathbf{x}\|^2 - \alpha) = \mathbf{x}^T(\mathbf{A} + \lambda^*\mathbf{I})\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c - \alpha\lambda^*. \quad (14)$$

- 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:

- If $\mathbf{x}^*$ is a minimizer of (TRS), then by the second order necessary conditions there exists $\lambda^* \geq 0$ such that

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

$$\|\mathbf{x}^*\|^2 \leq \alpha, \quad (16)$$

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

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

Proof Contd.

Necessity:

- If $\mathbf{x}^*$ is a minimizer of (TRS), then by the second order necessary conditions there exists $\lambda^* \geq 0$ such that

$$
\begin{aligned}
(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x}^* &= -\mathbf{b} && (15)\\
\|\mathbf{x}^*\|^2 &\leq \alpha, && (16)\\
\lambda^*(\|\mathbf{x}^*\|^2 - \alpha) &= 0, && (17)\\
\mathbf{d}^T(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{d} &\geq 0 \quad \text{for all } \mathbf{d} \text{ satisfying } \mathbf{d}^T\mathbf{x}^* = 0. && (18)
\end{aligned}
$$

- Need to show that (18) is true **for any d**.

Proof Contd.

Necessity:

- If $\mathbf{x}^*$ is a minimizer of (TRS), then by the second order necessary conditions there exists $\lambda^* \geq 0$ such that

$$\begin{aligned}
(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x}^* &= -\mathbf{b} && (15)\\
\|\mathbf{x}^*\|^2 &\leq \alpha, && (16)\\
\lambda^*(\|\mathbf{x}^*\|^2 - \alpha) &= 0, && (17)\\
\mathbf{d}^T(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{d} &\geq 0 \quad \text{for all } \mathbf{d} \text{ satisfying } \mathbf{d}^T\mathbf{x}^* = 0. && (18)
\end{aligned}$$

- Need to show that (18) is true **for any d**.
- Suppose on the contrary that there exists a **d** such that $\mathbf{d}^T\mathbf{x}^* > 0$ and $\mathbf{d}^T(\mathbf{A} + \lambda^* \mathbf{I})\mathbf{d} < 0$.

Proof Contd.

Necessity:

- If $\mathbf{x}^*$ is a minimizer of (TRS), then by the second order necessary conditions there exists $\lambda^* \geq 0$ such that

$$\begin{aligned} (\mathbf{A} + \lambda^*\mathbf{I})\mathbf{x}^* &= -\mathbf{b} && (15)\\ \|\mathbf{x}^*\|^2 &\leq \alpha, && (16)\\ \lambda^*(\|\mathbf{x}^*\|^2 - \alpha) &= 0, && (17)\\ \mathbf{d}^T(\mathbf{A} + \lambda^*\mathbf{I})\mathbf{d} &\geq 0 \quad \text{for all } \mathbf{d} \text{ satisfying } \mathbf{d}^T\mathbf{x}^* = 0. && (18) \end{aligned}$$

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

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

Proof Contd.

- In addition,

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

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

Total Least Squares

Consider the approximate set of linear equations:

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

Total Least Squares

Consider the approximate set of linear equations:

$$\mathbf{A}\mathbf{x} \approx \mathbf{b}$$

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

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

Total Least Squares

Consider the approximate set of linear equations:

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

- In the Least Squares (LS) approach we only assume that the RHS vector **b** is subjected to noise.

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

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

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

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

Eliminating the $\mathbf{E}$ and $\mathbf{w}$ variables

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

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

Eliminating the $\mathbf{E}$ and $\mathbf{w}$ variables

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

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

- The KKT conditions are necessary and sufficient for problem $(P_{\mathbf{x}})$.

Eliminating the $\mathbf{E}$ and $\mathbf{w}$ variables

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

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

- The KKT conditions are necessary and sufficient for problem $(P_{\mathbf{x}})$.
- Lagrangian: $L(\mathbf{E}, \mathbf{w}, \boldsymbol{\lambda}) = \|\mathbf{E}\|_F^2 + \|\mathbf{w}\|^2 + 2\boldsymbol{\lambda}^T[(\mathbf{A}+\mathbf{E})\mathbf{x} - \mathbf{b} - \mathbf{w}]$.

Eliminating the $\mathbf{E}$ and $\mathbf{w}$ variables

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

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

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

$$2\mathbf{E} + 2\boldsymbol{\lambda}\mathbf{x}^T = \mathbf{0} \qquad (\nabla_{\mathbf{E}} L = \mathbf{0}), \tag{19}$$

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

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

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

The New Formulation of (TLS)

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

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 $\mathbf{E} = -\frac{(\mathbf{A}\mathbf{x}-\mathbf{b})\mathbf{x}^T}{\|\mathbf{x}\|^2+1}$ and $\mathbf{w} = \frac{\mathbf{A}\mathbf{x}-\mathbf{b}}{\|\mathbf{x}\|^2+1}$

The New Formulation of (TLS)

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

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 $\mathbf{E} = -\frac{(\mathbf{A}\mathbf{x}-\mathbf{b})\mathbf{x}^T}{\|\mathbf{x}\|^2+1}$ and $\mathbf{w} = \frac{\mathbf{A}\mathbf{x}-\mathbf{b}}{\|\mathbf{x}\|^2+1}$

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

Solving the Fractional Quadratic Formulation

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

Solving the Fractional Quadratic Formulation

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

- (TLS') is the same as

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

Solving the Fractional Quadratic Formulation

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 $\mathbf{y} = \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}$):

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

where

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

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

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

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

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

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).

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

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

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.

- $f^* \geq g^*$.

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

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

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

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

Solving the Fractional Quadratic Formulation Contd.

We will consider the following relaxed version:

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

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

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

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

Main Result on TLS

Theorem. Assume that the following condition holds:

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

where

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

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 min. eigenvalue of $\mathbf{B}$.

Main Result on TLS

Theorem. Assume that the following condition holds:

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

where

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

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 min. eigenvalue of $\mathbf{B}$.

Proof.

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

Main Result on TLS

Theorem. Assume that the following condition holds:

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

where

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

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 min. 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

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

which is a contradiction to (24).