Optimality Conditions for constrained problems

Nonlinear Programming
Optimality Conditions
Unconstrained Case

$X^*$ is local min

$X^*$ is global min

Convex $f$

$\nabla f(x^*) = 0$

SOSC Hessian of $f$ pd

FONC

SONC Hessian of $f$ psd
Outline

Easy Case – Nonlinear objective with linear constraints
- Convex
- FONC
- Second Order conditions

Nonlinear objective with nonlinear constraints
- Convex
- FONC
- Second Order conditions
Assumptions

Objective and constraints are twice continuously differentiable for the purpose of this lecture.

Always take

\[ x \in X = \mathbb{R}^n \]
Easiest Problem

- Linear equality constraints

\[
\begin{align*}
\min f(x) & \quad f \in \mathbb{R}^n \\
\text{s.t. } Ax & \leq b & A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m \\
\end{align*}
\]

Constraints form a polyhedron
Only some constraints are active.

At $x^*$, only the constraint $A_1x \leq b_1$ is at bound/active.

At optimal solution, $x^*$, only one constraint is active.
Inequality Case

Inequality problem

$$x^* \in \arg \min f(x)$$

subject to

$$a_i x \leq b_i$$

$$i = 1 \ldots d$$

Linear Inequality KKT:

$$\nabla f(x^*) + \sum_i \lambda_i a_i = 0$$

$$Ax^* \leq b$$

$$\lambda_i (A_i x^* - b_i) = 0$$

$$\lambda \geq 0$$

Nonnegative Multipliers imply gradient points to the greater than Side of the constraint.
Complementarity Condition

Inequality FONC:

\[ \lambda_i (a_i x^* - b_i) = 0 \]

\[ \lambda_i > 0 \Rightarrow a_i x^* = b_i \quad \text{i.e. constraint is active} \]

\[ \lambda_i = 0 \Rightarrow \text{constraint may not matter. It can be dropped without changing result.} \]
Sufficient Conditions for Convex f with Linear Constraints

- $X^*$ minimizes $f$ over \( \{x|Ax \leq b\} \), where $f$ is convex, if and only if

\[
\Rightarrow \nabla f(x^*) + A' \lambda^* = 0
\]

\[
Ax^* \leq b
\]

\[
\lambda^* \geq 0
\]

\[
\lambda^*(Ax^* - b) = 0
\]

\[\text{KKT.}\]

You proved this in HW4
Alternative proof of KKT implies min

We know
\[ f(x^*) = L(x^*, u^*) = \min_x L(x, u^*) \]
since \( L(x, u^*) \) convex and \( \nabla L(x^*, u^*) = 0 \).

So
\[ f(x^*) = L(x^*, u^*) \leq L(x, u^*) \leq f(x) \quad \text{for all } x \ Ax \leq b. \]

so \( x^* \) is a global min.

Note that proof based on Lagrangian.
First Order Necessary Conditions for Linear inequality Constraints

If \( x^* \) is a local min of \( f \) over \( \{ x | Ax \leq b \} \), then for some vector \( \lambda^* \)

\[
\Rightarrow \nabla f(x^*) + A'\lambda^* = 0 \\
Ax^* \leq b \\
\lambda^* \geq 0 \\
\lambda^*(Ax^*-b) = 0
\]

\( KKT \)

You proved this in Homework 4
Sketch Proof

Every $d$ satisfying

$$\nabla f(x^*)'d < 0 \quad A_d \leq 0$$

is a feasible descent direction at $x^*$. The FONC make sure that this set is empty.

$\nabla f(x^*)'d < 0 \quad A_d \leq 0$ has no solution $d$ if and only if

$$-\nabla f(x^*) = A_d' \lambda \quad \lambda \geq 0$$

has a solution

Due to Farkas’s Lemma.
Second Order Necessary Conditions for Linear Constraints

If $x^*$ is a local min of $f$ over \{ $x | A x \leq b$ \}, and $Z$ is basis for null-space of active constraints then for some vector $\lambda^*$

\[
\nabla f(x^*) + A' \lambda^* = 0
\]

\[
A x^* \leq b
\]

\[
\lambda^* \geq 0
\]

\[
\lambda^*(A x^* - b) = 0
\]

and $Z' \nabla^2 f(x^*) Z$ is p.s.d.
Z for FONC

Z is null space matrix for matrix of active constraints at $x^*$

Index set of Active Constraints

$I = \{ i \mid a_i x^* = b_i \}$

Matrix of active constraints

$A_I$ is $A$ restricted to rows of $A$ in $I$

$Z = \text{null}(A_I)$
Proof by contradiction

\[ \exists p \text{ such that } pZ'\nabla^2 f(x^*)Zp < 0. \text{ Then let } d=Zp. \]

For \( \alpha > 0 \) sufficiently small, \( x^* + \alpha d \) is feasible for the inactive constraints and for the active constraints

\[ A_I(x^* + \alpha d) = A_I x^* + \alpha A_I Zp = b_I + 0. \]

Take a Taylor Series Expansion of \( L(x, \lambda^*) \) about \( x^* \),

\[ f(x^* + \alpha d) = L(x^* + \alpha d, \lambda^*) \]

\[ = L(x^*, \lambda^*) + \alpha \nabla L(x^*, \lambda^*) d + \alpha^2 d' \nabla^2 L(x^*, \lambda^*) d + \alpha^2 \| d \|^2 \ o(x^*, \alpha d) \]

So

\[ \frac{f(x^* + \alpha d) - f(x^*)}{\alpha^2} = d' \nabla^2 L(x^*, \lambda^*) d + \| d \|^2 \ o(x^*, \alpha d) \]

\[ f(x^* + \alpha d) - f(x^*) < 0, \text{ contradiction.} \]
Key Points

Note that proof used
Taylor Series Expansion of $L(x, \lambda^*)$ at $x^*$
It needs the Hessian of the Lagrangian to be p.s.d. with respect to potentially troublesome directions in the Tangent Cone

$$d^T \nabla^2 L(x^*, \lambda^*) d \geq 0 \quad \forall A_j d = 0$$
Second Order Sufficient Conditions for Linear Inequalities

If \((x^*, \lambda^*)\) satisfies

\[
Ax^* \leq b \quad \text{Primal feasibility}
\]
\[
\nabla f(x^*) + A' \lambda^* = 0 \quad \text{Dual feasibility}
\]
\[
\lambda^* \geq 0
\]
\[
\lambda^*(Ax^* - b) = 0 \quad \text{Complementarity}
\]

and SOSC \(Z_+ '\nabla^2 f(x^*)Z_+\) is p.d.

Then \(x^*\) is a strict local minimizer.
Sufficient Conditions for Linear Inequalities

where $Z_+$ is a basis matrix for $\text{Null}(A_+)$ and $A_+$ corresponds to nondegenerate active constraints

i.e.

$$A_+ = A_j$$

$$\{ j \mid A_j x^* = b_j, \lambda_j^* > 0 \}$$
Proof by contradiction sketch

The set of feasible directions at \( x^* \) or tangent cone is \( A_j d \leq 0 \). If \( d \) in the tangent cone satisfies \( a_i d < 0 \) for some \( i \) with \( \lambda_i > 0 \), then \( d \) is a direction of increase because

\[
f(x^* + \alpha d) = f(x^*) + \alpha \nabla f(x^*)' d + \alpha \| d \| o(x^*, \alpha d)
\]

\[
= f(x)^* - \alpha \lambda_j A_j' d + \alpha \| d \| o(x^*, \alpha d) \text{ where } \lambda_j A_j' d < 0
\]

So \( f(x^* + \alpha d) > f(x^*) \) for \( \alpha \) sufficiently small. So we need only be concerned about \( d \neq 0 \) such that \( A_j d = 0 \) or equivalently \( d = Z_+ p \) for some \( p \neq 0 \).
Proof by contradiction sketch
continued

For any \( d=Z_+p \neq 0 \),
take a Taylor Series Expansion of \( L(x, \lambda^*) \) about \( x^* \),
\[
f(x^*+\alpha d) = L(x^*+\alpha d, \lambda^*)
\]
\[
= L(x^*, \lambda^*) + \alpha \nabla L(x^*, \lambda^*)d + \alpha^2 \nabla^2 L(x^*, \lambda^*)d + \alpha^2 \| d \|^2 o(x^*, \alpha d)
\]
\[
= f(x^*) + \alpha^2 pZ'_+\nabla^2 f(x^*)Z'_+p + \alpha^2 \| d \|^2 o(x^*, \alpha d)
\]
where \( pZ'_+\nabla^2 f(x^*)Z'_+p > 0 \)
which implies \( f(x^*+\alpha d) - f(x^*) > 0 \) for \( \alpha \) sufficiently small,
so \( x^* \) is a strict local min.
Notes

- The proof is based on the Hessian of the Lagrangian function.
- Lagrangian must have positive curvature with respect to feasible directions.
- Other SOSC possible
  - e.g. Hessian psd in neighborhood of $x^*$
Sufficient Example

Find solution and verify SOSC

$$\begin{align*}
\min & \quad - \frac{1}{2} (x_1 + 1)^2 + \frac{1}{2} x_2^2 \\
\text{s.t.} & \quad 0 \leq x_1 \leq 1
\end{align*}$$

$$x^* = [1, 0]' \quad \lambda^* = [2, 0]$$
Linear Inequality Constraints - 1

\[
\begin{align*}
\text{min} & \quad -\frac{1}{2}(x^1 + 1)^2 + \frac{1}{2} x_2^2 \\
\text{s.t.} & \quad 0 \leq x_1 \leq 1
\end{align*}
\]

Put in standard form:

\[
\begin{align*}
\text{min} & \quad -\frac{1}{2}(x^1 + 1)^2 + \frac{1}{2} x_2^2 \\
\text{s.t.} & \quad x_1 - 1 \leq 0 \\
& \quad -x_1 \leq 0
\end{align*}
\]

\[x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\] by inspection

\[
\begin{align*}
\nabla f(x^*) &= \begin{bmatrix} -\frac{x_1^* + 1}{x_1^*} \\ x_2^* \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \\
\nabla^2 f(x^*) &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \\
\text{Active constraint is } x_1 \leq 1 \\
\therefore A_+ &= \begin{bmatrix} 1 & 0 \end{bmatrix}
\end{align*}
\]
Linear Inequality Constraints - II

\[ \nabla f(x^*) = \begin{bmatrix} - (x_1^* + 1) \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \]

\[ \nabla^2 f(x^*) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \]

Active constraint is \( x_1 \leq 1 \)

\[ \therefore A_+ = \begin{bmatrix} 1 & 0 \end{bmatrix} \]
Linear Inequality Constraints - III

KKT Conditions

\( A x^* = b \)

\( \nabla f(x^*) + A^T \lambda^* = 0 \)

\( \lambda^* \geq 0 \)

\( \lambda^* (Ax^* - b) = 0 \)

\( \lambda = \begin{bmatrix} ? \\ 0 \end{bmatrix} \) since second constraint is inactive

\( \nabla f(x^*) + A^T \lambda = 0 \Rightarrow \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \lambda_1 = 0 \Rightarrow \lambda_1 = 2 \geq 0 \)

KKT Point: \( x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \) \( \lambda^* = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \)
Linear Inequality Constraints - IV

Now look at SOSC

\[ A_+ = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad Z_+ = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

\[ \lambda^* = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \]

\[ \nabla^2 f(x^*) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \]

\[ Z_+ \nabla^2 f(x^*) = \begin{bmatrix} 1 \end{bmatrix} \text{ a p.d. matrix} \]

Therefore SOSC are satisfied.

\( X^* \) is a strict local minima
Why Necessary and Sufficient?

- **Sufficient conditions are good for?**
  - Way to confirm that a candidate point
  - is a minimum (local)
  - But...not every min satisfies any given SC

- **Necessary tells you:**
  - If necessary conditions don’t hold then you know you don’t have a minimum.
  - Under appropriate assumptions, every point that is a min satisfies the necessary cond.
  - Good stopping criteria
  - Algorithms look for points that satisfy Necessary conditions
Let’s Try

Solve the problem using above theorems: Verify $x^*=[8/9,4/18]$’ is optimal by checking KKT.

Also check SOSC

$$\begin{align*}
\min & \quad 1/2 x_1^2 + x_2^2 \\
\text{s.t.} & \quad 2x_1 + x_2 \geq 2 \\
& \quad x_1 - x_2 \leq 1 \\
& \quad x_1 \geq 0
\end{align*}$$
General Constraints

$$\min \quad f(x)$$

s.t. \quad h_i(x) = 0 \quad i \in E

\quad g_i(x) \leq 0 \quad i \in I

\quad x \in R^n
Active Set

The active set at a point $x$ consists of

- Equality constraints
- Inequality constraints at bound

$$A(x) = E \cup \{i \mid i \in I, g_i(x) = 0\}$$
For simplicity of notation

\[ \min f(x) \]

s.t. \[ g_i(x) \leq 0 \quad i \in I \]

\[ x \in \mathbb{R}^n \]

Focus on inequality case first
Lagrangian Function

Optimality conditions expressed using Lagrangian function

\[ L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) = f(x) + \lambda' g(x) \]

and Jacobian matrix

\[ \nabla g(x)' \]

were each row is a gradient of a constraint
Lagrangian Function

\[ L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) = f(x) + \lambda' g(x) \]

Lagrangian Gradient

\[ \nabla_x L(x, \lambda) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla_x g_i(x) \]

Lagrangian Hessian

\[ \nabla_{xx}^2 L(x, \lambda) = \nabla_{xx}^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla_{xx}^2 g_i(x) \]
Feasible descent directions

A point is not optimal at \( x \) if we can take a small step \( s \) that is feasible and decreases the function

\[
\min f(x_1, x_2) \quad s.t. \quad x_1^2 + x_2^2 - 1 \leq 0
\]
Feasible descent/Improvement directions

**Case 1: point in interior**

\[
\begin{align*}
\min & \quad f(x_1, x_2) \\
\text{s.t.} & \quad x_1^2 + x_2^2 \leq 1
\end{align*}
\]

Can only be optimal if \( \nabla f(x) = 0 \)
Feasible descent/Improvement directions

Case 2: point at bound

\[ \min f(x_1, x_2) \]
\[ s.t. \quad x_1^2 + x_2^2 \leq 1 \]

Any point in cone is an Improvement direction

If \( x^* \) is a local min then
\[ \nabla f(x)'s < 0 \quad \nabla g_1(x)'s < 0 \]

Must have no solution
Feasible descent/Improvement directions

\min f(x_1, x_2)
\text{s.t. } 1 - x_1^2 - x_2^2 \leq 0

\textbullet \ X^* \text{ is a min implies } \nabla f(x^*)'s < 0 \quad \nabla g_1(x^*)'s < 0 \text{ has no solution}

\textbullet \ \text{Same as saying there exists } u^* \text{ by T of Alt.}
\quad u_o \nabla f(x^*) + u^* \nabla g_1(x^*) = 0 \quad u_o, u^* \geq 0, \ [u_o, u^*] \neq 0

\textbullet \ \text{Same as saying there exists a Fritz John Point}
Fun with T of A

What is T of A for \( Ax < 0 \)?

\( Ax < 0 \) has a solution
\[ \iff \begin{cases} Ax \leq -\bar{1} \varepsilon \text{ has a solution} \\ \varepsilon > 0 \quad \text{here } \bar{1} \text{ is a vector of 1's} \end{cases} \]

\[ \iff \begin{bmatrix} A & \bar{1} \end{bmatrix} \begin{bmatrix} x \\ \varepsilon \end{bmatrix} \leq 0 \text{ has a solution} \]

\[ \iff \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ \varepsilon \end{bmatrix} > 0 \]

\[ \iff \begin{bmatrix} A' \end{bmatrix} u = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad u \geq 0 \quad \text{has no solution} \]

\[ \iff A'u = 0 \quad u \geq 0 \quad u \neq 0 \quad \text{has no solution} \]

Called Gordon’s T of A
Pseudo-normal implies KKT

For the Fritz John Point

\[ u_o \nabla f(x^*) + u^* \nabla g_1(x^*) = 0 \quad u_o, u^* \geq 0 \]

If the constraints are pseudo-normal then \( u_o > 0 \), so a KKT point exists

\[ g_1(x^*) \leq 0 \]
\[ \nabla f(x^*) + \lambda^* \nabla g_1(x^*) = 0 \quad \lambda^* \geq 0 \]
\[ \lambda^* g_1(x^*) = 0 \]
CQ imply Pseudo-normal

- The constraints are pseudo-normal at $x^*$ if a constraint qualification (CQ) is satisfied.

- Our favorite CQ
  - Linear - Constraints are linear
  - LICQ – gradients of the active constraints are linearly independent
  - Slater’s CQ – constraints are convex and there exists a strict interior point $g(x)<0$

See page 305-307 for more
First Order Necessary Conditions - Inequality

If $x^*$ is a local min of $f$ over $\{x|g(x)\leq 0\}$, and CQ satisfied then

\[
\begin{align*}
\text{there exists } \lambda^* &\geq 0 \\
g(x^*) &\leq 0 \\
\nabla_x L(x^*, \lambda^*) &= 0 \\
\lambda^* ' g(x^*) &= 0
\end{align*}
\]
FONC (KKT) are sufficient for convex programs.

If \((x^*, \lambda^*)\) satisfies FONC of convex program then \(x^*\) is a global minimum.

Proof of FONC implies global min

We know
\[
f(x^*) = L(x^*, \lambda^*) = \min_x L(x, \lambda^*)
\]

since \(L(x, \lambda^*)\) convex and \(\nabla L(x^*, \lambda^*) = 0\).

So
\[
f(x^*) = L(x^*, \lambda^*) \leq L(x, \lambda^*) = f(x) \quad \text{for all } x \quad g(x) \leq 0.
\]

so \(x^*\) is a global min.
If \((x^*, \lambda^*)\) satisfies

\[ g(x^*) \geq 0 \quad \text{Primal feasibility} \]

\[ \nabla f(x^*) + \nabla g(x^*) \lambda^* = 0 \]

(equivalently \(\nabla_x L(x^*, \lambda^*) = 0\))

\[ \lambda^* \geq 0 \quad \text{(for inequalities only)} \]

\[ \lambda^* g(x^*) = 0 \quad \text{Complementarity} \]

and \(\text{SOSC } Z_+ \nabla^2_{xx} L(x^*) Z_+ \text{ is p.d.} \)

Then \(x^*\) is a strict local minimizer
Second Order Sufficient Conditions Nonlinear Inequality
where \( Z_+ \) is a basis matrix for \( \text{Null}(A_+) \) and \( A_+ \) corresponds to Jacobian of nondegenerate active constraints)

i.e.

For the jth row of Jacobian \( \Rightarrow \)

\[ g_j(x^*) = 0 \quad \text{Active Constraint} \]

\[ \lambda_j^* > 0 \quad \text{Nondegenerate} \]

\[ (A_+)_j = \nabla g_j(x^*)' \]
Sufficient Example

Find solution and verify SOSC

\[
\begin{align*}
\min & \quad -\frac{1}{2}(x_1 + 1)^2 - \frac{1}{2}x_2^2 \\
\text{s.t.} & \quad \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \leq \frac{1}{2}
\end{align*}
\]

\[x^* = [1, 0]' \quad \lambda^* = 2\]
Nonlinear Inequality Constraints - I

\[
\begin{align*}
\min & \quad -\frac{1}{2}(x_1 + 1)^2 - \frac{1}{2}x_2^2 \\
\text{s.t.} & \quad \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \leq \frac{1}{2}
\end{align*}
\]

Guess \( x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \)

\[
\nabla f(x) = \begin{bmatrix} -(x_1 + 1) \\ -x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}
\]

\[
\nabla^2 f(x^*) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}
\]
Nonlinear Inequality Constraints - II

Has one active constraint: \( \frac{1}{2} x_1^2 + \frac{1}{2} x_2^2 \leq \frac{1}{2} \)

Jacobian: \( \nabla g(x)^T = [x_1, x_2] = [1, 0] \)

\[ L(x, \lambda) = f(x) + \lambda' g(x) \]

\[ \nabla_x L(x, \lambda) = \nabla f(x) + \lambda_1 \nabla g_1(x) \]

\[ = \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]
Nonlinear Inequality Constraints - III

\[ L(x, \lambda) = f(x) + \lambda' g(x) \]

\[ \nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x) + \sum_i \lambda_i \nabla^2 g_i(x) \]

\[ \nabla_{xx}^2 L(x^*, \lambda^*) = \begin{bmatrix}
-1 & 0 \\
0 & -1 
\end{bmatrix} + 2 \begin{bmatrix}
1 & 0 \\
0 & 1 
\end{bmatrix} \]

= \begin{bmatrix}
1 & 0 \\
0 & 1 
\end{bmatrix} \quad \leftarrow \text{positive definite} \]

\[ Z_+ = \begin{bmatrix}
0 \\
1 
\end{bmatrix}, \text{ since } \nabla g(x)^T = [1 \ 0] \]

\[ Z_+^T \nabla_{xx}^2 L(x, \lambda) Z_+ = \begin{bmatrix}
0 & 1 \\
0 & 1 
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
0 & 1 
\end{bmatrix} \begin{bmatrix}
0 \\
1 
\end{bmatrix} = 1 \quad \leftarrow \text{positive definite} \]

so SOSC satisfied, \( x^* \) is a strict local minimum
Sufficient Example

Find solution and verify SOSC

\[
\begin{align*}
\min & \quad x_2 \\
\text{s.t.} & \quad x_1^2 - x_2 = 0
\end{align*}
\]

\[x^* = [0, 0]' \quad \lambda^* = 1\]
Nonlinear Inequality Constraints - V

\[ \text{min } x_2 \]
\[ \text{s.t. } x_1^2 - x_2 = 0 \]
\[ x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ by observation} \]

\[ L(x, \lambda) = f(x) + \lambda^T g(x) \]
\[ = x_2 + \lambda (x_1^2 - x_2) \]
\[ \nabla_x L(x, \lambda) = \nabla f(x) + \lambda \nabla g(x) \]
\[ = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 2x_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

\[ \nabla_x L(x^*, \lambda^*) = 0 \Rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda^* \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0 \Rightarrow \lambda^* = 1 \]
Nonlinear Inequality Constraints - VI

\[ \nabla g(x)^T = \begin{bmatrix} 0 & -1 \end{bmatrix} \quad \therefore \quad Z_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

\[ \nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x) + \sum_i \lambda_i \nabla^2 g_i(x) \]

\[ = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + (1) \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \]

\[ = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \]

\[ Z_+^T \nabla_{xx}^2 L(x, \lambda) Z_+ = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \rightarrow \text{positive definite} \]

So SOSC are satisfied, and \( x^* \) is a strict local minimum.
Second Order Necessary Conditions Inequality

If $x^*$ is a local min of $f$ over $\{x|g(x) \leq 0\}$, $Z$ is a null-space matrix of the Jacobian $\nabla g(x^*)'$ of the active constraints, and CQ satisfied at $x^*$ then

there exists $\lambda^*$

$\nabla_x L(x^*, \lambda^*) = 0$ or equivalently $Z'\nabla f(x^*) = 0$

$g(x^*) \leq 0$

$\lambda^*' g(x^*) = 0$

and $Z'\nabla^2_{xx} L(x^*) Z$ is p.s.d.
Necessary Example

Show optimal solution $x^*=[1,0]'$ is pseudonormal (LICQ satisfied called regular in many text) and find KKT point

$$\begin{align*}
\text{max} & \quad x_1 \\
\text{s.t.} & \quad x_1^2 + x_2^2 \leq 1 \\
& \quad (x_1 - 1)^3 - x_2 \leq 0
\end{align*}$$

KKT point

$$\begin{align*}
x^* &= [1, 0]' \\
\lambda^* &= [1/2, 0]'
\end{align*}$$
FONC

For general problems with CQ at $x^*$, we know $x^*$ is a local min implies then there exists a corresponding KKT point.

- Fritz John Point always exists
- CQ ensures pseudo-normality
- Pseudo-normality implies KKT exists.
General SOSC

Consider problem min of $f$ over
\[ \{x | g(x) \leq 0, h(x) = 0\}, \]

If there exists $x^*, \lambda^*_I \geq 0, \lambda^*_E$ such that
\[
0 = \nabla_x L(x^*, \lambda^*_I, \lambda^*_E) = \nabla f(x^*) + \sum_{i \in E} \lambda^*_i \nabla g_i(x^*) + \sum_{i \in I} \lambda^*_i \nabla h_i(x^*)
\]
\[
g(x^*) \leq 0, \ h(x^*) = 0
\]
\[
\lambda^* g(x^*) = 0 \quad \text{and} \quad Z^+ \nabla^2_x L(x^*)Z^+ \text{ is p..d}
\]
where $Z^+$ is the null space matrix for $A_+(x^*)$ (active constraints with non-zero multipliers)

then $x^*$ is a strict local minimizer
Other SOSC

- KKT conditions are sufficient for any convex program. Same proof above works.
- Alternative SOSC, Hessian of Lagrangian is psd for directions that matter in a neighborhood of the solution $x^*$. See other NLP texts such as Bazarra and Sherali.
Necessary Conditions for General

If $x^*$ satisfies CQ and is a local min of $f$ over $\{x|g(x)\leq 0, h(x)=0\}$, there exists $\lambda^*_I \geq 0, \lambda^*_E$

$$0 = \nabla_x L(x^*, \lambda^*_I, \lambda^*_E) = \nabla f(x^*) + \sum_{i \in E} \lambda^*_i \nabla g_i(x^*) + \sum_{i \in I} \lambda^*_i \nabla h_i(x^*)$$

$g(x^*) \leq 0, h(x^*) = 0$

$\lambda^* g(x^*) = 0$

and $Z' \nabla^2_{xx} L(x^*) Z$ is p.s.d

where $Z$ is the null space matrix for $A(x^*)$ (active constraints)
KKT Summary

- X* is local min
- KKT Satisfied
- X* is global min
- Convex f
- Convex constraints
- CQ
- SOSC $Z_+ \nabla^2 L(x^*, \lambda^*) Z_+ \text{ p.d.}$
- SONC $Z \nabla^2 L(x^*, \lambda^*) Z \text{ p.s.d.}$