Math 164 Notes

Brendan Connelly

September to December 2025

1 Preliminary Lectures

Definition (Quadratic Form)**.** Let $A \in \mathbb{R}^{n\times n}$. A *quadratic form* is a function

$$f : \mathbb{R}^n \to \mathbb{R}, \qquad f(x) = x^\top Ax.$$

We will assume A is symmetric. Indeed, if $f(x) = x^\top Bx$ for some (possibly non-symmetric) B, then there exists a symmetric matrix A such that $f(x) = x^\top Ax$, namely

$$A = \frac{B + B^\top}{2}.$$

Hence, without loss of generality, we may take A to be symmetric.

Example. $f(x_1, x_2) = x_1^2 + x_2^2$ is positive definite.

Definition (Definiteness of a Quadratic Form)**.** Let $A = A^\top \in \mathbb{R}^{n\times n}$ and define

$$f(x) = x^\top Ax.$$

Then the quadratic form f is called:

- **Positive definite (PD)** if $\forall x \neq 0,\ f(x) > 0$.
- **Positive semidefinite (PSD)** if $\forall x \in \mathbb{R}^n,\ f(x) \geq 0$.

  Example: $f(x_1, x_2) = x_1^2$.
- **Negative definite (ND)** if $\forall x \neq 0,\ f(x) < 0$.
- **Negative semidefinite (NSD)** if $\forall x \in \mathbb{R}^n,\ f(x) \leq 0$.
- **Indefinite** if there exist x_1, x_2 such that $f(x_1) > 0$ and $f(x_2) < 0$.

Lemma. Let $A = A^\top$ and $f(x) = x^\top Ax$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. Then:

- f is **positive definite (PD)** iff all $\lambda_i > 0$.
- f is **negative definite (ND)** iff all $\lambda_i < 0$.
- f is **positive semidefinite (PSD)** iff all $\lambda_i \geq 0$.
- f is **negative semidefinite (NSD)** iff all $\lambda_i \leq 0$.

Lemma (Sylvester's Criterion)**.** **Note:** This criterion does not provide conclusions about semi-definiteness.
Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad A = A^\top.$$

Consider the **leading principal minors:**

$$M_1 = |a_{11}|, \quad M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad M_3 = |A|, \quad \text{and so on.}$$

Lemma (continued). Assume $M_1, M_2, \ldots, M_n \neq 0$. Then:

- If $M_1 > 0, M_2 > 0, \ldots, M_n > 0$, then A is **positive definite**.
- If $M_1 < 0, M_2 > 0, M_3 < 0, \ldots$ (signs alternate), then A is **negative definite**.
- Otherwise, A is **indefinite**.

Definition (Hessian)**.** Let $f \in C^2(\mathbb{R}^n)$. The **Hessian** of f at x is the $n \times n$ matrix

$$\nabla^2 f(x) := \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}.$$

Definition (Jacobian)**.** Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable map. The **Jacobian matrix** of g at x is

$$Dg(x) = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(x) & \cdots & \dfrac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(x) & \cdots & \dfrac{\partial g_m}{\partial x_n}(x) \end{pmatrix}.$$

In particular, for $f : \mathbb{R}^n \to \mathbb{R}$, the gradient is given by

$$\nabla f(x) = [Df(x)]^\top.$$

Theorem (Chain Rule)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^r$ and $g : \mathbb{R}^k \to \mathbb{R}^n$ be C^1 functions, and define

$$h = f \circ g : \mathbb{R}^k \to \mathbb{R}^r.$$

Then $h \in C^1(\mathbb{R}^k)$ and

$$Dh(x) = Df(g(x))\, Dg(x).$$

Theorem (First-Order Taylor Approximation)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function and fix $x \in \mathbb{R}^n$. Then, for all $y \in \mathbb{R}^n$, we have

$$f(y) = f(x) + \nabla f(x)^\top (y - x) + R(x, y - x),$$

where

$$\lim_{y \to x} \frac{R(x, y - x)}{\|y - x\|} = 0.$$

The linear term

$$T_1(y) = f(x) + \nabla f(x)^\top (y - x)$$

is called the **first-order Taylor approximation** of f at x.

Theorem (Second-Order Taylor Approximation)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function and fix $x \in \mathbb{R}^n$. Then, for all $y \in \mathbb{R}^n$, we have

$$f(y) = f(x) + \nabla f(x)^\top (y - x) + \tfrac{1}{2}(y - x)^\top \nabla^2 f(x)(y - x) + R_2(x, y - x),$$

where

$$\lim_{y \to x} \frac{R_2(x, y - x)}{\|y - x\|^2} = 0.$$

The quadratic term

$$T_2(y) = f(x) + \nabla f(x)^\top (y - x) + \tfrac{1}{2}(y - x)^\top \nabla^2 f(x)(y - x)$$

is called the **second-order Taylor approximation** of f at x.

Definition (Convex Set)**.** A set $\Omega \subseteq \mathbb{R}^n$ is called *convex* if for all $x, y \in \Omega$ and all $\alpha \in [0, 1]$,

$$\alpha x + (1 - \alpha) y \in \Omega.$$

Definition (Convex Function)**.** Let $\Omega \subseteq \mathbb{R}^n$ be convex. A function $f : \Omega \to \mathbb{R}$ is called *convex* if for all $x, y \in \Omega$ and all $\alpha \in [0, 1]$,

$$f(\alpha x + (1 - \alpha) y) \le \alpha f(x) + (1 - \alpha) f(y).$$

Lemma. Let $\Omega \subseteq \mathbb{R}^n$ be open and convex, and let $f \in C^2(\Omega)$.

- f is **convex** $\iff$ for all $x \in \Omega$, the Hessian $\nabla^2 f(x)$ is **positive semidefinite (PSD)**.
- f is **concave** $\iff$ for all $x \in \Omega$, the Hessian $\nabla^2 f(x)$ is **negative semidefinite (NSD)**.

Definition. Let $\Omega \subseteq \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$.

- A point $x^* \in \Omega$ is called a **global minimum** of f on Ω if

$$f(x^*) \le f(y) \quad \forall\, y \in \Omega.$$

- A point $x^* \in \Omega$ is called a **local minimum** of f on Ω if there exists $\varepsilon > 0$ such that

$$f(x^*) \le f(y) \quad \forall\, y \in \Omega \text{ with } \|y - x^*\| < \varepsilon.$$

Definition (Feasible direction)**.** Let $\Omega \subseteq \mathbb{R}^n$ and consider a point $x^* \in \Omega$. A direction $d \in \mathbb{R}^n \setminus \{0\}$ is called a **feasible direction** of x^* (with respect to Ω) if there exists $\bar{\alpha} > 0$ such that

$$x^* + \alpha d \in \Omega \quad \forall \alpha \in (0, \bar{\alpha}].$$

Remark. If d is a feasible direction at x^*, then for all $\beta > 0$, βd is also a feasible direction.

Theorem (First Order Necessary Condition (FONC))**.** Let $f \in C^1(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is open. If $x^* \in \Omega$ is a local minimum of f over Ω, then for every feasible direction d at x^*, it holds that

$$d^\top \nabla f(x^*) \geq 0.$$

Proof. Let d be a feasible direction, and without loss of generality assume $\|d\| = 1$. By the first-order Taylor expansion of f at $x = x^*$, evaluated at $y = x^* + \alpha d$ for $\alpha > 0$, we have

$$f(x^* + \alpha d) = f(x^*) + \nabla f(x^*)^\top (\alpha d) + R(x^*, \alpha d),$$

where

$$\lim_{\|\alpha d\| \to 0} \frac{R(x^*, \alpha d)}{\|\alpha d\|} = 0 \quad \text{or equivalently} \quad \lim_{\alpha \to 0} \frac{R(x^*, \alpha d)}{\alpha} = 0.$$

□

Corollary (FONC for Interior Points)**.** If Ω is open or $x^* \in \operatorname{int}(\Omega)$, then

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

Theorem (SONC)**.** Let $f \in C^2(\Omega)$, and let $x^* \in \Omega$ be a local minimum of f over Ω. Then, for every feasible direction d at x^*, it holds that

$$d^\top \nabla f(x^*) = 0 \quad \Longrightarrow \quad d^\top \nabla^2 f(x^*) d \geq 0.$$

Proof. Without loss of generality, let $\|d\| = 1$ and d be a feasible direction such that $d^\top \nabla f(x^*) = 0$.

By the second-order Taylor expansion,

$$f(x^* + \alpha d) = f(x^*) + \nabla f(x^*)^\top (\alpha d) + \tfrac{1}{2} (\alpha d)^\top \nabla^2 f(x^*) (\alpha d) + R_2(x^*, \alpha d),$$

where $\lim_{\alpha \to 0} \dfrac{R_2(x^*, \alpha d)}{\alpha^2} = 0$.

Since x^* is a local minimum and $f(x^* + \alpha d) \geq f(x^*)$ for small $\alpha > 0$, it follows that $d^\top \nabla^2 f(x^*) d \geq 0$. □

Theorem (SOSC)**.** Consider $f \in C^2(\Omega)$ and $x^* \in \Omega$ being an interior point. If it holds that

$$\nabla f(x^*) = 0 \quad \text{and} \quad \nabla^2 f(x^*) \text{ is positive definite (PD)},$$

then we can conclude that x^* is a **strict local minimum** of f over Ω.

Proof. First, recall the following lemma:

Let $H = H^\top \in \mathbb{R}^{n \times n}$ be positive definite. Then there exists $\lambda > 0$ such that

$$y^\top H y \geq \lambda \|y\|^2, \quad \forall y \in \mathbb{R}^n,$$

where λ is the smallest eigenvalue of H.

—

Now, applying the second-order Taylor expansion around x^*:

$$f(x^* + h) = f(x^*) + \tfrac{1}{2} h^\top \nabla^2 f(x^*) h + R_2(x^*, h),$$

where $\lim\limits_{h \to 0} \dfrac{R_2(x^*, h)}{\|h\|^2} = 0$.

Since $\nabla^2 f(x^*)$ is positive definite, we have

$$h^\top \nabla^2 f(x^*) h \geq \lambda \|h\|^2$$

for some $\lambda > 0$. Therefore, for sufficiently small h,

$$f(x^* + h) - f(x^*) \geq \tfrac{1}{2} \lambda \|h\|^2 + R_2(x^*, h) > 0.$$

Hence, x^* is a strict local minimum. $\square$

Problem.

$$\min_{x \in \Omega} f(x)$$

where

$$\Omega \subseteq \mathbb{R}^n, \quad f : \Omega \to \mathbb{R}$$

and f is a convex function.

Theorem. Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and $f : \Omega \to \mathbb{R}$ be a convex function. Then:

1. Every **local minimum** of f is also a **global minimum**.

$$x^* \text{ local min} \implies x^* \text{ global min}.$$

2. If f is **strictly convex**, then there exists **at most one global minimum**.

$$f \text{ strictly convex} \implies \text{unique global minimum}.$$

(e.g., consider $f(x) = e^x$.)

3. If Ω is open and $f \in C^1(\Omega)$, then

$$x^* \in \Omega \text{ is a global minimum} \iff \nabla f(x^*) = 0.$$

Remark. Practice problem: Prove part (3). It could be on the exam!

Problem.

$$\max_{x \in \Omega} f(x)$$

where

$$\Omega \subseteq \mathbb{R}^n, \quad f : \Omega \to \mathbb{R}$$

and f is a concave function.

Theorem. Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and $f : \Omega \to \mathbb{R}$ be a concave function. Then:

1. Every **local maximum** of f is also a **global maximum**.
$$x^* \text{ local max} \implies x^* \text{ global max.}$$
2. If f is **strictly concave**, then there exists **at most one global maximum**.
$$f \text{ strictly concave} \implies \text{unique global maximum.}$$
(e.g., consider $f(x) = e^x$.)
3. If Ω is open and $f \in C^1(\Omega)$, then
$$x^* \in \Omega \text{ is a global maximum} \iff \nabla f(x^*) = 0.$$

Theorem (Weierstrass)**.** Let $\Omega \subseteq \mathbb{R}^n$ be non-empty and compact, and let $f : \Omega \to \mathbb{R}$ be continuous. Then:

1. f attains at least one **global minimizer** over Ω.
2. The set of all global minimizers of f is compact.

(The same holds for global maximizers.)

Also known as the Extreme Value Theorem.

Lecture 6

Proposition. Let $\Omega \subseteq \mathbb{R}^n$ be convex and let $f : \Omega \to \mathbb{R}$ be strictly concave. Then f has **at most one global maximum** over Ω.

Proof. Since f is strictly concave, for all $\alpha \in (0, 1)$ and all distinct $x, y \in \Omega$,
$$f(\alpha x + (1 - \alpha)y) > \alpha f(x) + (1 - \alpha)f(y).$$
Suppose, for contradiction, that $x^*, y^* \in \Omega$ are distinct global maximizers. Then $f(x^*) = f(y^*) = \max_\Omega f$. By strict concavity,
$$f(\alpha x^* + (1 - \alpha)y^*) > \alpha f(x^*) + (1 - \alpha)f(y^*) = f(x^*),$$
contradicting maximality of x^*. Hence, there can be at most one global maximizer. ☐

Proposition. Let $f : \mathbb{R}^n \to \mathbb{R}$ and let $\Omega \subseteq \widetilde{\Omega} \subseteq \mathbb{R}^n$. **Claim (True or False):** If $x^* \in \Omega$ is a local minimizer of f over Ω, then x^* is also a local minimizer of f over $\widetilde{\Omega}$.

Counterexample. Let $f(x) = x$, $\Omega = [0, \infty)$, and $\widetilde{\Omega} = \mathbb{R}$. Then $x^* = 0$ is a local minimum of f over Ω, but not a local minimum over $\widetilde{\Omega}$.

Definition (Strictly Unimodal)**.** Function $f : [a_0, b_0] \to \mathbb{R}$ is strictly unimodal on its domain iff there exists some $x_0 \in [a_0, b_0]$ such that for all $x_1, x_2 \in [a_0, b_0]$ with either
$$x_0 < x_1 < x_2 \quad x_0 > x_1 > x_2$$
such that
$$f(x_0) < f(x_1) < f(x_2)$$
Note: this does not imply convexity/concavity

Definition (Possibly Correct alternative - Strictly Unimodal)**.** A function $f : [a_0, b_0] \to \mathbb{R}$ is *strictly unimodal* on its domain if there exists a point $x_0 \in (a_0, b_0)$ such that:

$$\begin{cases} f(x_1) < f(x_2), & \text{for all } a_0 \le x_1 < x_2 \le x_0, \\ f(x_1) > f(x_2), & \text{for all } x_0 \le x_1 < x_2 \le b_0. \end{cases}$$

That is, f is strictly increasing on $[a_0, x_0]$ and strictly decreasing on $[x_0, b_0]$.

Concept (Zero-order search methods)**.** Goal:

- Start with $[a_0, b_0]$ and in one iteration, shorten it
- Only use values of f, i.e., zero-order information

Idea:

- Start with $[a_0, b_0]$
- Choose points $a_0, a_1 < b_1 < b_0$ and evaluate $f(a_1), f(b_1)$
- If $f(a_1) > f(b_1)$, the local minimum cannot lie in $[a_0, a_1)$ and we can continue with search interval $[a_1, b_0]$.

Likewise, if $f(a_1) < f(b_1)$,the local minimum cannot lie in $(b_1, b_0]$ and we can continue with search interval $[a_0, b_1]$.

There is something to prove to show that the local minimum cannot be in the excluded intervals

Lecture 7

Concept (Search once given interval)**.** We want to find the min $f : \mathbb{R} \to \mathbb{R}$ over $[a_0, b_0]$. Assumption is that f is unimodal and we are given this search interval. We want to continue shortening the intervals

Concept (Minimal evaluations of $f(\cdot)$ and iterations - efforts reduce numerical complexity)**.** We want to have as few evaluations of $f(\cdot)$ as possible. In intervals, we will reuse test points.

We have some options:

1. split search interval in same proportions ($\rho : (1 - \rho)$, make the length of the interval after one iteration as small as possible
2. Reuse test points but allow for different splitting ratios in each iteration:

$$\rho_1, \dots, \rho_N$$

   and specify N iterations a priori

Theorem (Golden Section Search)**.** We order $a_0 < a_1 < b_1 < b_0$. WLOG $|b_0 - a_0| > 1$. Split in proportion given by $\rho \in (0, \frac{1}{2})$. So,

$$a_1 - a_0 = b_0 - b_1 \propto \rho \qquad b_1 - a_0 = b_0 - a_1 \propto 1 - \rho$$

We obtain the golden ratio since we want

$$\frac{1-\rho}{1} = \frac{\rho}{1-\rho}$$

Hence,

$$\rho = \frac{3-\sqrt{5}}{2} = 1 - \frac{1}{\phi}$$

Iteration of Golden section search $[a_0, b_0]$:

$$a_1 = a_0 + \rho(b_0 - a_0) \qquad b_1 = a_0 + (1-\rho)(b_0 - a_0)$$

Example (Golden Section Search Two Iterations)**.** Consider $f(x) = 2x^3 - 3x^2$ on $[0, 2]$. This turns out to be strictly unimodal on the interval. Now for iteration #1:

$$a_1 = a_0 + \rho(b_0 - a_0) = 0 + \rho(2) = 0.7639$$

$$b_1 = a_0 + (1-\rho)(b_0 - a_0) = 0 + 2(1-\rho) = 1.2361$$

Now, we compare $f(a_1) = -0.8591 < -0.8065 = f(b_1)$. So, we consider $[a_0, b_1]$.

Remark (Tolerance & Iteration Count)**.** Given a tolerance ε, how many iterations will I need to perform?

We start with length $b_0 - a_0$. Then, our length continuously decreases by a factor of $\frac{1}{\phi} = 1 - \rho$. After N iterations, the interval will be

$$(b_0 - a_0) \times (1-\rho)^N$$

So, we are solving

$$(b_0 - a_0) \times (1-\rho)^N < \varepsilon \qquad \implies \qquad \frac{\log(\frac{\varepsilon}{b_0 - a_0})}{\log(1-\rho)}$$

Theorem (Fibonacci Method - Reuse test points & Adapt interval length method)**.** Reuse test points but allow for different splitting ratios in each iteration:

$$\rho_1, \ldots, \rho_N$$

and specify N iterations a priori

We aim for the shortest possible interval after N iterations.

WLOG, intervals of length 1. We need to choose
$a_k - a_{k-1} = b_{k-1} - b_k = \rho_k \implies b_k - a_k = 1 - 2\rho_k$
In the new iteration,

$$\rho_{k+1}(1-\rho_k) = a_{k+1} - a_k \qquad \rho_{k+1}(1-2\rho_k) = b_k - b_{k+1}$$

So then,

$$1 - 2\rho_k = \rho_{k+1}(1-\rho_k) \implies \rho_{k+1} = \frac{1-2\rho_k}{1-\rho_k}$$

After N iterations, (WLOG $b_0 - a_0 = 1$), the length of the interval is

$$\min \rho_k \forall k \prod_k (1-\rho_k)$$

that satisfies $\rho_k \in [0, \frac{1}{2}]$ and $\rho = \frac{1-2\rho_k}{1-\rho_k}$.

If you solve this,

$$p_1 = 1 - \frac{F_N}{F_N+1}, \qquad p_2 = 1 - \frac{F_{N-1}}{F_N}, \ldots$$

We define the fibonnaci numbers such that

$$F_0 = F_1 = 1 \qquad F_{k+1} = F_k + F_{k-1}$$

Example (Fibonacci Method Iterations & Issue)**.** Iteration # 1 on $[a_0, b_0]$. We then set

$$a_1 = a_0 + \rho(b_0 - a_0) \qquad b_1 = a_0 + (1-\rho)$$

But we have an issue:
Suppose $N = 3$. Then,

$$\rho_1 = 1 - \tfrac{3}{5} \quad \rho_2 = 1 - \tfrac{2}{3} \quad \rho_3 = 1 - \tfrac{1}{2}$$

And, we can see that the last iteration is impossible. So, as an Ad Hoc solution, we move last two points ε apart.

$$\tilde{\rho} = \tfrac{1}{2} - \varepsilon$$

Lecture 8

Concept (Zero Order methods for minimization of functions $f : \mathbb{R} \to \mathbb{R}$)**.** Assumptions:

1. f is strictly unimodal on $\mathbb{R}$. *Note: this is a very strong assumption.*

Question: how do we find this interval? I.e., how to enclose (bracket) our local min into an interval $[a, b]$?
Zero Order Bracketing

- Start with choosing $x_0 < x_1 < x_2$. If $f(x_0) > f(x_1) > f(x_2)$, then, keep selecting new test points $x_0, x_1, x_2, \ldots$, until $f(x_{k-1}) > f(x_k) < f(x_{k+1})$
- if $f(x_0) < f(x_1) < f(x_2)$, then start indexing the other way.

Recommend: $|x_{k+1} - x_k| = 2|x_k - x_{k-1}|$

Concept (First Order methods for minimization of functions $f : \mathbb{R} \to \mathbb{R}$)**.** Suppose we again have a search interval $[a, b]$. If I can find $a < b$ with $f'(a) < 0$ and $f'(b) > 0$, then the local minimum lies in $[a, b]$.

To find this interval, choose $x_0 \in \mathbb{R}$. Then, if $f'(x_0) < 0$, choose $x_0 < x_1 < x_2 < \ldots$ until $f'(x_k) > 0$. Reverse if $f'(x_0) > 0$.

Example (Bisection Method - First Order minimization)**.** Bisection Method - 1 iteration on $[a, b]$. Compute $f'(\frac{a+b}{2})$. Then

- if $f'(\frac{a+b}{2}) > 0$, choose $[a, \frac{a+b}{2}]$.
- if $f'(\frac{a+b}{2}) < 0$, choose $[\frac{a+b}{2}, b]$.

Note: Tolerance after N iterations is $\frac{|b-a|}{2^N}$

Proposition (Newton's Method)**.** *Main Idea:*

- Start with x_0
- At each iteration k with x_k approximate at f in the neighborhood of x_k by the best possible quadratic approximation. (second order Taylor)

$$T_2(y) = f(x_k) + f'(x_k)(y - x_k) + \tfrac{1}{2} f''(x_k)(y - x_k)^2$$

- Find the stationary point of T_2 and use it as the next iterate x_{k+1}.

$$T_2'(x) = f'(x_k) + f''(x_k)(y - x_k) = 0 \implies \boxed{y = x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}}$$

- This method is very fast. While this will work very well for a convex function, but not a concave function (just looking for stationary) so could find a maximizer. So, the method strongly depends on where we start.

  This is looking for a root of the first order derivative

Proposition (Newton's Method for Root Finding)**.** Given $g : \mathbb{R} \to \mathbb{R}$. We want to find x such that $g(x) = 0$. Let's use a linear approximation.

- Start with $x_0 \in \mathbb{R}$
- In iter k with x_k, approximate g by linear function around x_k. Then,

$$T_1(y) = g(x_k) + g'(x_k)(y - x_k)$$

- Choose the root of $T_1(\cdot)$ as the next iterate x_{k+1}

$$\implies \boxed{x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}}$$

Lecture 9

Proposition (Secant Method - Newton's Method with approximation of second derivative)**.** Recall

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

We consider x_{k-1}. We can then approximate our second derivative via

$$f''(x_k) \approx \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}$$

Hence, the secant method produces

$$x_{k+1} = x_k - \frac{f'(x_k)}{\frac{f'(x_k)-f'(x_{k-1})}{x_k - x_{k-1}}} = x_k - \frac{x_k - x_{k-1}}{f'(x_k) - f'(x_{k-1})} f'(x_k)$$

We should initialize with x_0, x_{-1}. When to stop?

- When $|f'(x_k)| < \varepsilon$
- When $\frac{|f(x_{k+1}) - f(x_k)|}{\max\{1, |f(x_k)|\}} < \varepsilon$
- When $\frac{|x_{k+1} - x_k|}{\max\{1, x_k|\}} < \varepsilon$

Done with chapter on single variable - move to multiple variable case

---

Concept (Chapter 5 - Gradient Methods)**.** Numerical methods for $\min f : \mathbb{R}^n \to \mathbb{R}$ over open $\Omega \in \mathbb{R}^n$. We will assume $f \in C^1(\Omega)$.

Iterative Methods: construct a sequence $x_0, x_1, x_2, \ldots$ with the aim of $(x_k)_{k \in \mathbb{N}}$ converging to a stationary point. This is best possible hope or we need much stronger assumptions.

Descent Methods: $f(x_{k+1}) < f(x_k)$

Point-direction-step size scheme: $x_{k+1} = x_k + \alpha_k d_k$. We will focus on these as classical numerical methods are based on this.

1. How to choose a direction?
2. How to choose a step size?

Definition (Descent Direction)**.** Fix $f \in C^1, x \in \mathbb{R}^n$. A (feasible) direction $d \in \mathbb{R}^n \setminus \{0\}$ is a descent direction at $x \in \mathbb{R}^n$ iff there exists some $\bar{\alpha} > 0$ such that $\forall \alpha \in (0, \bar{\alpha}),\ f(x + \alpha d) < f(x)$.

Note: If d is descent direction, then $\beta * d$ is descent direction for $\beta > 0$.

Lemma (Sufficient conditions for descent direction)**.** Fix $f \in C^1, x \in \mathbb{R}^n$. A direction $d \in \mathbb{R}^n \setminus \{0\}$ satisfying $d^T \nabla f(x) < 0$ is a descent direction at x.

Proof. WLOG, let d be normalized with $d^T \nabla f(x) < 0$. Then, T_y first order Taylor expansion at x, which we will evaluate at $x + \alpha d$. Then

$$f(x + \alpha d) = f(x) + \alpha \nabla f(x)^T d + R(x, \alpha d)$$

with $\lim_{\alpha \to 0} \frac{R(x, \alpha d)}{\alpha} = 0$. For $\alpha > 0$,

$$\frac{f(x + \alpha d) - f(x)}{\alpha} = d^T \nabla f(x) + \frac{R(x, \alpha d)}{\alpha}$$

For small enough $\alpha > 0$, the RHS will be dominated by $d^T \nabla f(x)$, the negative term, so the RHS is strictly negative. Hence, the LHS is too, so

$$\frac{f(x + \alpha d) - f(x)}{\alpha} < 0 \implies f(x + \alpha d) < f(x)$$

□

Remark (The reverse lemma is not true)**.** d descent direction at x does not imply $d^T \nabla f(x) < 0$. Take a local maximizer with gradient 0. Then, any d will decrease yet $d^T \nabla f(x) = 0$.

Corollary. Let $f \in C^1, x \in \mathbb{R}^n$ is not a stationary point of f. Then,

- $d = -\nabla f(x)$
- $d = -D\nabla f(x)$ for $D = D^T$ a positive definite matrix

Lecture 10

Concept (Premise of day)**.** Numerical methods for $\min f : \mathbb{R}^n \to \mathbb{R}$ over open $\Omega \in \mathbb{R}^n$. We will assume $f \in C^1(\Omega)$.

Iterative Methods: construct a sequence $x_0, x_1, x_2, \ldots$ with the aim of $(x_k)_{k\in\mathbb{N}}$ converging to a stationary point. This is best possible hope or we need much stronger assumptions. We will study the **Point-direction-step size scheme** where $x_{k+1} = x_k + \alpha_k d_k$. We will focus on these as classical numerical methods are based on this.

We will focus on how to choose a step size. Assuming: the current iterate and (descent) direction d_k are fixed. This now becomes

$$\Phi_k(\alpha) = f(x_k + \alpha d_k) \quad \Phi : \mathbb{R} \to \mathbb{R}$$

Taking derivatives produces

$$\Phi_k'(\alpha) = Df(x_k + \alpha d_k)Dg(\alpha) = \nabla f(x_k + \alpha d_k)^T d_k$$

We can see that

$$\Phi_k'(0) = \nabla f(x_k)^T d_k$$

Concept (Option 1 to choose step size - Line Search)**.**

$$\min_{\alpha \geq 0} \Phi_k(\alpha) = \min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

In most cases, needs to be solved numerically, i.e., secant method, Newton's method, etc. Each iteration will then take relatively long. General observation, not the best of computational time.

Example. Let's consider an example in an exam situation: $f(x, y) = x^2 + y^2 - xy$. Given initial guess $(x_0, y_0) = (0, 0)$ with $d_k = -\nabla f(x_k, y_k)$ and optimal step size.

$$\nabla f(x, y) = \begin{pmatrix} 2x - y - 1 \\ 2y - x + 2 \end{pmatrix} \implies d_0 = -\nabla f(0, 0) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

So,

$$\Phi_0(\alpha) = f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ -2 \end{pmatrix}\right) = f(\alpha_1 - 2\alpha) = \alpha^2 + 4\alpha^2 + 2\alpha^2 - \alpha - 4\alpha = 7\alpha^2 - 5\alpha$$

Hence,

$$\Phi_0' = 12\alpha - 5 = 0 \implies \alpha_0 = \tfrac{5}{14}$$

Thus,

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \tfrac{5}{14} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

In general, would also be good idea to check second order condition.
EXPECT TO DO THIS ON EXAM

Concept (Option 2 to choose step size - Predetermined step size)**.**

- e.g. fixed step size $\alpha \equiv \frac{1}{2}$
- trouble is, how do we choose?
- at very least, make sure satisfying descent condition, i.e., $f_{k+1} < f(x_k)$ - obviously does not guarentee convergence of method

Example (Failure of option 2)**.**

1. Too large step size

   Consider $f(x) = x^2, x_0 = 2, d_{2l} = -1, d_{2l+1} = 1$. Step size is $\alpha_k = 2 + 3\frac{1}{2^{k+1}}$.

   We start with intial guess two. Then, we end up at $x_1 = -1.5$, $x_2 = 1.25$, both converge to $1, -1$, not 0. See lecture notes for derivation.

$$x_{2l} = 1 + \frac{1}{2^{2l}} \qquad x_{2l+1} = -1 - \frac{1}{2^{2l+1}}$$

2. Too small step size

   Consider $f(x) = x^2, x_0 = 2, d_k = -1, \alpha = \frac{1}{2^{k+1}}$. Sequence converges to 1, not zero. In lecture notes, it derives

$$x_{k+1} = 2 - \left(1 - \frac{1}{2^{k+2}}\right) = 1 + 2^{-k-2}$$

Concept (Option 3 - Low Accuracy Line Search)**.**

1. Start solving $\min_{\alpha \geq 0} \Phi_k(\alpha)$ (e.g. via secant method) but terminate once (some) condition verify you found step size that is neither too small nor too large

Armigo Conditions: Let $\varepsilon \in (0, 1), j > 1$. Then, to check α_k not too large

$$\Phi_k(\alpha_k) \leq \Phi_k(0) + \varepsilon \alpha_k \Phi_k'(0)$$

and to check α_k not too small

$$\Phi_k(j\alpha_k) \geq \Phi_k(0) + \varepsilon j \alpha_k \Phi_k'(0)$$

This won't be tested on exam, just that this exists.

Lecture 11

Concept (Today's topics)**. Gradient Methods**: terminology is not necessarily fixed

- direction $d_k = -\nabla f(x_k)$
- step size α_k, various
- Steepest descent method (optimal step size)

$$x_{k+1} = x_k - \alpha_k \nabla f(x) \qquad \alpha_k = \arg\min_{\alpha > 0} f(x_k - \alpha \nabla f(x))$$

Lemma. Assuming $\nabla f(x_k) \neq 0$, we have that $d_k = -\nabla f(x_k)$ is a descent direction for f at x_0. Recall:

We will focus on how to choose a step size. Assuming: the current iterate and (descent) direction d_k are fixed. This now becomes

$$\Phi_k(\alpha) = f(x_k + \alpha d_k) \quad \Phi : \mathbb{R} \to \mathbb{R}$$

Taking derivatives produces

$$\Phi_k'(\alpha) = Df(x_k + \alpha d_k)Dg(\alpha) = \nabla f(x_k + \alpha d_k)^T d_k$$

We can see that

$$\Phi_k'(0) = \nabla f(x_k)^T d_k$$

Proposition (Locally steepest direction - Why ∇f is optimal)**.**

$$\Phi_{dk}'(0) = \nabla f(x_k)^T d_k$$

tells us the slope of Φ_{dk}.

Fix a function f and a point x (should be your current iterate) that is not a stationary point. Among all descent directions d at a point x that are normalized (i.e. $\|d\| = 1$), the choice of $\overline{d} = \frac{-\nabla f(x)}{\|\nabla f(x)\|}$ minimizes the slope of the map

$$\alpha \mapsto \Phi_d(\alpha) = f(x + \alpha d) \quad \text{at} \quad \alpha = 0$$

Note: this is only true locally.

Proof. By Cauchy Schwartz, we have that

$$\left| \frac{d}{d\alpha} f(x + \alpha d) \right|_{\alpha=0} = \left| \nabla f(x)^T d \right| = |\langle \nabla f(x), d \rangle| \leq \|\nabla f(x)\| \cdot \|d\|$$

For descent direction $d^T \nabla f(x) \leq 0$ and normalized d, we have that

$$d^T \nabla f(x) = -|d^T \nabla f(x)| \geq -\|\nabla f(x)\| \|d\| = -\|\nabla f(x)\|$$

For $\overline{d} = \frac{-\nabla f(x)}{\|\nabla f(x)\|}$, we obtain

$$\overline{d}^T \nabla f(x) = -\frac{\nabla f(x)^T \nabla f(x)}{\|\nabla f(x)\|} = -\|\nabla f(x)\|$$

Since we obtained our lower bound, we are done.

□

Proposition (Optimal step size for quadratic f)**.** Let $f(x) = \frac{1}{2} x^\top Q x + b^\top x + c$ with $Q = Q^\top$. In the steepest descent method $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$, the optimal step size α_k that minimizes $f(x_k - \alpha \nabla f(x_k))$ is given by

$$\boxed{\alpha_k = \frac{\nabla f(x_k)^\top \nabla f(x_k)}{\nabla f(x_k)^\top Q \nabla f(x_k)}}.$$

Proposition. Let $\{x_0, x_1, x_2, \ldots\}$ be a sequence generated by the steepest descent method, and x_k is not a stationary point. Then, $f(x_{k+1}) < f(x_k)$.

Proof. We know that $d_k = -\nabla f(x)$ is a descent direction, i.e., there exists a $\overline{\alpha} > 0$ such that for all $\alpha \in (0, \overline{\alpha}), f(x_k + \alpha d_k) < f(x_k)$. We know that α_k is optimal. Therefore,

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) \leq f(x_k + \tfrac{\overline{\alpha}}{2} d_k) < f(x_k)$$

$\square$

Proposition. Let $\{x_0, x_1, x_2, \ldots\}$ be a sequence generated by the steepest descent method, and x_k is not a stationary point. Then,

$$\langle x_{k+1} - x_k, x_{k+2} - x_{k+1} \rangle = 0$$

Proof. Recall $x_{k+1} = x_k - \alpha_k \nabla f(x) \qquad \alpha_k = \arg\min_{\alpha > 0} f(x_k - \alpha \nabla f(x))$. So,

$$\langle x_{k+1} - x_k, x_{k+2} - x_{k+1} \rangle = \alpha_k \alpha_{k+1} \langle \nabla f(x_k), \nabla f(x_{k+1}) \rangle$$

Since α_k solve minimization on open set, it must satisfy FONC,

$$0 = \Phi_k'(\alpha_k) = \nabla f(x_k + \alpha_k d_k)^T d_k = -\nabla f(x_{k+1})^T \nabla f(x_k)$$

So the above term is zero, showing orthogonality. $\square$

Concept (Summary of current situation)**.** A very important question in optimization is the question of convergent. If I use a given method to generate a sequence, do I have any guarantees that the sequence will converge and how fast? For us, the intuitive answer is that the gradient based method under reasonably weak methods do converge, but they converge slowly (zig zagging path, not straight)

Lecture 13

Concept (Newton's Method, chapter 6)**.** For $f : \mathbb{R} \to \mathbb{R}$, recall

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

Now, for $\mathbb{R}^n \to \mathbb{R}$, assume $f \in C^2(\mathbb{R}^n)$. Idea: Locally (around current iterate x^k), aproximate f by quadratic. Find stationary points of quadratic approximation.

$$q(y) = f(x_k) + \nabla f(x_k)^T (y - x_k) + \tfrac{1}{2}(y - x_k)^T \nabla^2 f(x_k)(y - x_k)$$

Stationary point of $q : \nabla q(y) = 0$. Then,

$$0 = \nabla q(y) = \nabla f(x_k) + \nabla^2 f(x_k)(y - x_k)$$

So, $\nabla^2 f(x_k)(y - x_k) = -\nabla f(x_k)$. Hence, assuming invertibility,

$$y = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

This gives our iteration for the pure newton's method.

Properties

- direction $d_k = -\left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$
- step size $\alpha_k \equiv 1$

Possible Problems

- Let's assume d_k is a descent direction. But, the pure Newton's method is using a step size of 1. α_k could be too large.
- d_k might not be a descent direction when $\nabla^2 f$ is not positive definite, similar to the case when we may accidentally be choosing an ascent direction in the one-dimensional case.
- Evaluating the Hessian matrix can be very costly computationally. Furthermore, evaluating inverse of $\nabla^2 f$, or more commonly, solving linear system for d_k.

We will discuss solutions to each of these issues

Proposition (Modified Newton Method)**.** Let us assume the hessian matrix is positive definite. Hence, d_k is a descent direction. To avoid this problem, don't insist on a fixed step size. Do optimal step size of limited line search.

Lemma (How to force symmetric matrix to be positive definite)**.** Consider $A = A^T$. Then, there exists a $\mu > 0$ such that $A + \mu I$ is positive definite.

Proof. (Sketch) Symmetric A gives real values eigenvalues possibly with multiplicity $\lambda_1, \lambda_2, \ldots, \lambda_n$ which we will assume are in increasing order, with eigenvectors $v_1, \ldots, v_n$. Then,

$$(A + \mu I)v_i = Av_i + \mu I v_i = (\lambda + \mu)v_i$$

$\square$

Proposition (Levenberg-Marquardt modification)**.**

$$x_{k+1} = x_k - \alpha_k(\nabla^2 f(x_k) + \mu I)^{-1}\nabla f(x_k)$$

For appropriately chosen $\mu \in \mathbb{R}$, d_k is a descent direction. We can also choose optimal step size. This solves problems 1 and 2.

Proposition (Quasi Newton's Method)**.** We will try to approximate the second derivative.

Lecture 14

Concept (Classes of Methods we have focused on in the last two weeks)**.** For the gradient methods, we can look at steepest descent.

$$x_{k+1} = x_k - \alpha_k \nabla f(x) \qquad \alpha_k = \arg\min_{\alpha > 0} f(x_k - \alpha \nabla f(x))$$

In contrast, for the pure Newton's Method, it is also an iterative method. We are not just using the negative gradient but the negative gradient multiplied by inverse of hessian with constant step size of 1.

$$y = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

The steepest descent method is arguably best for one step. But, with Newton's method, we are constantly replacing an objective function with a quadratic approximation and looking for a stationary point of that approximation. In terms of computationally expensiveness, the direction for steepest descent is cheap because it only requires evaluating the gradient. It is slightly more expensive in step size aspect because we require solving a single variable minimization problem in choosing the step size. We have to fully perform a line search in each iteration. In contrast, for the Newton's method, the step size is very cheap, it is fixed. But, to compute the direction, we need to evaluate the hessian and then solve a linear system (inverse of matrix).

In terms of guarantees of convergence, weaker assumptions needed to converge. It converges linearly, much more slowly. For the Newton Method, we need stronger convergence. It converges much faster, quadratic even.

Concept (Conjugate Gradient Methods)**. The Advantages**

- Only evaluates $\nabla f(\cdot)$
- Quadratic f: Converges in N iterations

Assume we are given some $Q = Q^T \in \mathbb{R}^{N \times N}$. We can then define

Definition (Q conjugate)**.** Assume we are given some $Q = Q^T \in \mathbb{R}^{N \times N}$. Directions $d_0, d_1, \ldots, d_k \in \mathbb{R}^N$ are Q conjugate iff $d_j^T Q d_i = 0$ for $i \neq j$. If $Q = I$, reduces to orthogonality. You can define a more general Q inner product in this way. It will only be a true inner product if Q is positive definite.

Lemma. Let $Q = Q^T$ be positive (negative) definite. If $d_0, d_1, \ldots, d_k \in \mathbb{R}^N$ are non-zero and Q conjugate, they will be linearly independent.

Example (Conjugate gradient for toy case)**.** Suppose $f(x) = \frac{1}{2}x^T Q x - b^T x$. Assumptions:

- $Q = Q^T$ is positive definite (so convex)
- We are given N $Q-$conjugate directions $d_0, d_1, \ldots, d_{N-1} \in \mathbb{R}^N$

Toy algorithm

$$x^{k+1} = x^k + \alpha^k d^k \qquad \alpha^k = -\frac{\nabla f(x^k)^T d^k}{(d^k)^T Q d^k}$$

You can prove that if you do this you will get an algorithm that will do rather well if you start with arbitrary directions, will converge in N iterations. You can also show that you will be obtaining gradients that will be orthogonal to all the previous directions. We are interested in this second observation because we would like to translate this toy algorithm to a more practical algorithm.

A variant is generating conjugate directions throughout the algorithm

$$d_0 = -\nabla f(x_0) \qquad d_k = -\nabla f(x_k) + \beta_{k-1} d_{k-1}$$

Proposition (Conjugate Gradient Algorithms for General Function f)**.**

$$d_0 = -\nabla f(x_0) \qquad d_k = -\nabla f(x_k) + \beta_{k-1} d_{k-1}$$

If someone gives you a β,

$$x_{k+1} = x_k + \alpha_k d_k \qquad \alpha_k = \arg\min_{\alpha \in \mathbb{R}} f(x_k + \alpha d_k)$$

There are a number of formulas, one is called the **Fletcher-Reeves Formula**

$$\beta_{k-1} = \frac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_{k-1})^T \nabla f(x_{k-1})}$$

2 Lecture 15

Concept (Modern Methods - Quasi Newton Methods)**.** We remember that Newton's Method had 3 issues

$$y = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

- Let's assume d_k is a descent direction. But, the pure Newton's method is using a step size of 1. α_k could be too large.
- d_k might not be a descent direction when $\nabla^2 f$ is not positive definite, similar to the case when we may accidentally be choosing an ascent direction in the one-dimensional case.
- Evaluating the Hessian matrix can be very costly computationally. Furthermore, evaluating inverse of $\nabla^2 f$, or more commonly, solving linear system for d_k.

We discussed solutions to the first two problems. For the final, we can consider Quasi Newton's Methods. We want to approximate $(\nabla^2 f(\cdot))^{-1}$ by some matrix.

- Do it iteratively
- make the approx matrix sym and PSD by construction
- start with $H^0 = I$
- For $k = 0, 1, 2, \ldots$

  $$H^{k+1} = H^k + \Delta H^k$$

  with ΔH^k built from $\nabla f(x^k)$. We then have $x^{k+1} = x^k - \alpha^k H^k \cdot \nabla f(x^k)$, and we know this is a descent direction.

The most simple formula for the updating matrix is the rank one formula, where we abbreviate $\nabla x^k = \alpha^k d^k = x^{k+1} - x^k$ and $\nabla g^k = \nabla f(x^{k+1}) - \nabla f(x^k)$

$$\nabla H^{k+1} = \frac{(\Delta x^k - H^k \Delta g^k)(\Delta x^k - H^k \Delta g^k)^T}{(\Delta g^k)^T (\Delta x^k - H^k \Delta g^k)}$$

Concept (Stochastic Gradient Methods)**.** Idea for a neural network. You have your data $\{x_i\}_{i \in [N]}$. For each data point, $x_i \mapsto$ some loss function $Q_i(\theta)$. We are usually trying to solve the problem of minimizing our average loss, i.e.,

$$\min_\theta \frac{1}{N} \sum_{i=1}^{N} Q_i(\theta)$$

If we tried to use standard gradient methods, we would need

$$\nabla f(\theta) = \frac{1}{N} \sum_{i=1}^{N} \nabla Q_i(\theta)$$

for every single loss function, for every single point.

The main idea of stochastic gradient methods is to approximate $\nabla f(\theta)$ through a few points in the sample.

Midterm

3 Lecture 16

Concept (Intro to constrained optimization)**.** We want to solve for $X \subset \mathbb{R}^n$, which we call our problem (P)

$$\min_{x \in X} f(x) \qquad h_i(x) = 0 \qquad \forall i = 1, \dots m \quad \text{and} \quad g_i(x) \leq 0 \quad \forall i = 1, \dots r$$

We write $h : X \to \mathbb{R}^m$ and $g : X \to \mathbb{R}^r$

$$h(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{pmatrix} = 0 \qquad g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} \leq 0$$

We still need to consider our feasible set:

$$\Omega = \{x \in X \mid h(x) = 0 \quad g(x) \leq 0\}$$

Definition (Optimal Value)**.**

$$f* := \inf_{x \in X} f(x)$$

Definition (Optimal Solution)**.** Global Minimizer of f over Ω.

4 Lecture 17

Concept (Lagrange Multipliers and variables)**.** Useful for many applications. Main idea is that it allows us to transform a constrained optimization problem to an unconstrained one. We introduce m variables λ_i for equality constraints, and μ_i for the inequality constraints. Then,

$$L : X \times \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R} \qquad L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{i=1}^{r} \lambda_i g_i(x)$$

We can also write

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$

Definition (Lagrange Multipliers)**.** We normally refer to Lagrange multipliers as specific values of Lagrange variables. The point of $\lambda* \in \mathbb{R}^m$ and $\mu* \in \mathbb{R}^r$ are called Lagrange Multipliers for (P) iff

$$\mu* \geq 0 \quad \text{and} \quad f* = \inf_{x \in X} L(x, \lambda*, \mu*)$$

The rough idea of why this is useful is that once you find the Lagrange multipliers, you can just optimize the unconstrained $L(x, \lambda*, \mu*)$.

Remark. The form of the constraint function matters and the lagrange multiplier doesn't have to exist.

Lemma. Let $x* \in \Omega$ be an optimal solution of (P) and $(\lambda*, \mu*)$ be a Lagrange multiplier for (P). Then,

$$f(x*) = L(x*, \lambda*, \mu*) \quad \text{and} \quad \mu_j * \cdot g_j(x*) = 0 \quad j = 1, \dots, r$$

Proof.

$$f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \sum_{i=1}^{r} \lambda_i^* g_i^*(x^*) \leq f(x^*)$$

□

by $(\lambda^*, \mu*)$ Lagrange Mult.

$$f(x^*) = f^* = \inf_{x \in X} L(x, \lambda^*, \mu^*) \leq L(x^*, \lambda^*, \mu^*)$$

So,

$$0 = \sum \mu_j^* g_j(x^*) \implies \mu_j^* g_j(x^*) = 0$$

Theorem (Saddle Point Theorem)**.** Point $x^* \in \Omega$ is an optimal solution of (P) and (λ^*, μ^*) is a Lagrange multiplier for (P) if and only if for all $x \in X$ and for all $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^r$

$$L(x^*, \lambda, \mu) \leq L(x^*, \lambda^*, \mu^*) \leq L(x, \lambda^*, \mu^*).$$

Lecture 18

Concept (Equality Constrained Problems)**.** We are considering $\min_{x \in \mathbb{R}^n} f(x)$ such that $h(x) = 0$, which we call problem (P).

Definition. A feasible point $\overline{x} \in \Omega$ is regular for (P) iff $\{\nabla h_1(\overline{x}), \dots, h_m(\overline{x})\}$ are linearly independent.

Proposition (First Order Necessary Condition (FONC))**.** Assume $f, h_1, \dots, h_m \in C^1$. If $x^* \in \Omega$ is a regular local minimizer of (P), then there exists a unique vector $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T \in \mathbb{R}^m$ such that

$$\nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) = 0 \quad \text{thus} \quad \nabla L(x^*, \lambda^*) = 0$$

It is important to be careful to consider possible irregular feasible points in addition to the general case.

Lecture 19

Definition. An inequality constraint $g_j(x) \leq 0$ is said to be *active* at x^* if

$$g_j(x^*) = 0.$$

Let $x^* \in \Omega$. The set of active constraints at x^* is

$$J(x^*) = \{\, j : g_j(x^*) = 0 \,\}.$$

A feasible point $x^* \in \Omega$ is called *regular* if the gradients

$$\nabla h_i(x^*),\ \nabla g_j(x^*) \qquad i = 1, \ldots, m,\ j \in J(x^*)$$

are linearly independent.

Lecture 20

Remark. The conditions $\mu^* \geq 0$ and $(\mu^*)^\top g(x^*) = 0$ are known as the *complementary slackness conditions.*

Lecture 21

Concept. Recall

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{such that} \quad h(x) = 0, g(x) \leq 0 \qquad (P)$$

Observation: If $x^* \in \Omega$ is a (regular) local min. of (P), then x^* is a (regular) loc. min. of

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{such that} \quad h(x) = 0, g_j(x) = 0 \quad j \in J(x^*) \qquad (\tilde{P})$$

Theorem (FONC for (P) - Karush-Kuhn-Tucker Conditions)**.** If x^* is a regular local min of (P), then there exists λ^*, μ^* such that

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) + \sum \mu_j^* \nabla g_j(x^*) = 0$$

with $\mu_j^* \geq 0$ and $\mu_j^* g_j(x^*) = 0 \quad j = 1, \ldots, r$

Theorem (Second Order Necessary Condition (SONC) for $\tilde{P}$)**.** Assume that $f, h_1, \ldots, h_m, g_1, \ldots, g_r \in C^2$. Suppose $x^* \in \Omega$ is a regular local minimum of problem $(\tilde{P})$:

$$\min f(x) \qquad \text{s.t. } h(x) = 0,\ g_j(x) = 0\ \ \forall j \in J(x^*).$$

Then there exist unique multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu_j^* \in \mathbb{R}$ for all $j \in J(x^*)$ such that the first-order condition holds:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j \in J(x^*)} \mu_j^* \nabla g_j(x^*) = 0. \tag{FONC–$\tilde{P}$}$$

Define the subspace of first–order feasible variations:

$$V(x^*) = \left\{ y \in \mathbb{R}^n : y^\top \nabla h_i(x^*) = 0\ \forall i = 1, \ldots, m, \quad y^\top \nabla g_j(x^*) = 0\ \forall j \in J(x^*) \right\}.$$

Then for all $y \in V(x^*)$ it holds that

$$y^\top \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j \in J(x^*)} \mu_j^* \nabla^2 g_j(x^*) \right) y \ \geq 0. \tag{SONC–$\tilde{P}$}$$

Remark. The space $V(x^*)$ is the subspace of first–order feasible directions for $(\tilde{P})$, i.e., the directions in which all equality constraints (the h_i and the active g_j) remain satisfied to first order.

Theorem (Second–Order Necessary Conditions (SONC))**.** Let $f, h, g \in C^2$ and let a regular point $x^* \in \Omega$ be a local minimizer of (P). Then there exist multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^r$ such that

$$\mu^* \geq 0, \qquad (\mu^*)^\top g(x^*) = 0,$$

and the first–order condition holds:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j^* \nabla g_j(x^*) = 0.$$

Define

$$V(x^*) = \left\{ y \in \mathbb{R}^n : y^\top \nabla h_i(x^*) = 0 \ (i = 1, \dots, m),\ y^\top \nabla g_j(x^*) = 0 \ (j \in J(x^*)) \right\}.$$

Then for all $y \in V(x^*)$ it holds that

$$y^\top \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j \in J(x^*)} \mu_j^* \nabla^2 g_j(x^*) \right) y \ \geq \ 0.$$

Theorem (Sufficient Condition)**.** Let $f, h, g \in C^2$. Suppose there exist $x^* \in \Omega$, multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^r$ such that

$$\mu^* \geq 0, \qquad (\mu^*)^\top g(x^*) = 0,$$

and

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j^* \nabla g_j(x^*) = 0.$$

Define

$$V(x^*, \mu^*) = \left\{ y \in \mathbb{R}^n : y^\top \nabla h_i(x^*) = 0 \ (i = 1, \dots, m), \quad y^\top \nabla g_j(x^*) = 0 \ (j \in J(x^*, \mu^*)) \right\},$$

where

$$J(x^*, \mu^*) = \{ \, j : g_j(x^*) = 0,\ \mu_j^* > 0 \, \}.$$

If for all $y \in V(x^*, \mu^*) \setminus \{0\}$ it holds that

$$y^\top \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j=1}^{r} \mu_j^* \nabla^2 g_j(x^*) \right) y > 0,$$

then x^* is a *strict local minimizer* of (P).

Theorem (Kuhn-Tucker Theorem)**.** Consider the problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{such that} \quad g(x) \leq 0 \qquad (CP)$$

where f, g_i are C^1 and convex. If there exists $x^* \in \Omega$ and $\mu^* \in \mathbb{R}^r$ such that

$$\nabla_x L(x^*, \mu^*) = 0 \quad \text{and} \quad \mu^* > 0, (\mu^*)^T g(x^*) = 0$$

then x^* is an optimal solution (global min) of (CP).