APM462: Nonlinear Optimization

Tianyu Du

November 2, 2019

Contents

1 Preliminaries

1.1 Mean Value Theorems and Taylor Approximations.

Definition 1.1. Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$, the **gradient** of f at $x \in S$, if exists, is a vector $\nabla f(x) \in \mathbb{R}^n$ characterized by the property

$$\lim_{v \to 0} \frac{f(x+v) - f(x) - \nabla f(x) \cdot v}{||v||} = 0 \tag{1.1}$$

Theorem 1.1 (The First Order Mean Value Theorem)**.** Let $f \in C^1(\mathbb{R}^n, \mathbb{R})$, then for any $x, v \in \mathbb{R}^n$, there exists some $\theta \in (0, 1)$ such that

$$f(x+v) = f(x) + \nabla f(x + \theta v) \cdot v \tag{1.2}$$

Proof. Let $x, v \in \mathbb{R}^n$, define $g(t) := f(x + tv) \in C^1(\mathbb{R}, \mathbb{R})$.
By the mean value theorem on $\mathbb{R}^{\mathbb{R}}$, there exists $\theta \in (0, 1)$ such that $g(0 + 1) = g(0) + g'(\theta)(1 - 0)$, that is, $f(x + v) = f(x) + g'(\theta)$. Note that $g'(\theta) = \nabla f(x + \theta v) \cdot v$. ■

Proposition 1.1 (The First Order Taylor Approximation)**.** Let $f \in C^1(\mathbb{R}^n, \mathbb{R})$, then

$$f(x+v) = f(x) + \nabla f(x) \cdot v + o(||v||) \tag{1.3}$$

that is

$$\lim_{||v|| \to 0} \frac{f(x+v) - f(x) - \nabla f(x) \cdot v}{||v||} = 0 \tag{1.4}$$

Proof. By the mean value theorem, $\exists \theta \in (0, 1)$ such that $f(x + v) - f(x) = \nabla f(x + \theta v) \cdot v$.
The limit becomes $\lim_{||v|| \to 0} \frac{[\nabla f(x+\theta v) - \nabla f(x)] \cdot v}{||v||} = \lim_{||v|| \to 0; x + \theta v \to x} \frac{[\nabla f(x+\theta v) - \nabla f(x)] \cdot v}{||v||}$.
Since $f \in C^1$, $\lim_{x+\theta v \to x} \nabla f(x + \theta v) = \nabla f(x)$.
And $\frac{v}{||v||}$ is a unit vector, and every component of it is bounded, as the result, the limit of inner product vanishes instead of explodes. ■

Theorem 1.2 (The Second Order Mean Value Theorem)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, then for any $x, v \in \mathbb{R}^n$, there exists $\theta \in (0, 1)$ satisfying

$$f(x+v) = f(x) + \nabla f(x) \cdot v + \frac{1}{2} v^T \nabla^2 f(x + \theta v) v \tag{1.5}$$

Proposition 1.2 (The Second Order Taylor Approximation)**.** Let $f : C^2(\mathbb{R}^n, \mathbb{R})$ function, and $x, v \in \mathbb{R}^n$, then

$$f(x+v) = f(x) + \nabla f(x) \cdot v + \frac{1}{2} v^T \nabla^2 f(x + \theta v) v + o(||v||^2) \tag{1.6}$$

that is

$$\lim_{||v|| \to 0} \frac{f(x+v) - f(x) - \nabla f(x) \cdot v - \frac{1}{2} v^T \nabla^2 f(x) v}{||v||^2} = 0 \tag{1.7}$$

Proof. By the second mean value theorem, there exists $\theta \in (0, 1)$ such that the limit is equivalent to

$$\lim_{||v|| \to 0} \frac{1}{2} \left(\frac{v}{||v||} \right)^T \left[\nabla^2 f(x + \theta v) - \nabla^2 f(x) \right] \frac{v}{||v||} \tag{1.8}$$

Since $f \in C^2$, the limit of $[H_f(x + \theta v) - H_f(x)]$ is in fact $\mathbf{0}_{n \times n}$. And every component of unit vector $\frac{v}{||v||}$ is bounded, the quadratic form converges to zero as an immediate result. ■

It is often noted that the gradient at a particular $x_0 \in dom(f) \subseteq \mathbb{R}^n$ gives the direction f increases most rapidly. Let $x_0 \in dom(f)$, and v be a unit vector representing a *feasible direction* of change. That is, there exists $\delta > 0$ such that $x_0 + tv \in dom(f)$ $\forall t \in [0, \delta)$. Then the rate of change of f along feasible direction v can be written as

$$\left.\frac{d}{dt}\right|_{t=0} f(x_0 + tv) = \nabla f(x_0) \cdot v = ||\nabla f(x_0)||\ ||v|| \cos(\theta) \tag{1.9}$$

where $\theta = \angle(v, \nabla f(x_0)$. And the derivative is maximized when $\theta = 0$, that is, when v and ∇f point the same direction.

1.2 Implicit Function Theorem

Theorem 1.3 (Implicit Function Theorem)**.** Let $f : C^1(\mathbb{R}^{n+1}, \mathbb{R})$, let $(a, b) \in \mathbb{R}^n \times \mathbb{R}$ such that $f(a, b) = 0$. If $\nabla f(a, b) \neq 0$, then $\{(x, y) \in \mathbb{R}^n \times \mathbb{R} : f(x, y) = 0\}$ is locally a graph of a function $g : \mathbb{R}^n \to \mathbb{R}$.

Remark 1.1. $\nabla f(x_0) \perp$ level set of f near x_0.

2 Convexity

2.1 Terminologies

Definition 2.1. Set $\Omega \subseteq \mathbb{R}^n$ is **convex** if and only if

$$\forall x_1, x_2 \in \Omega,\ \ \lambda \in [0, 1],\ \ \lambda x_1 + (1 - \lambda)x_2 \in \Omega \tag{2.1}$$

Definition 2.2. A function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex** if and only if Ω is convex, and

$$\forall x_1, x_2 \in \Omega,\ \ \lambda \in [0, 1],\ \ f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{2.2}$$

Definition 2.3. A function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is **strictly convex** if and only if Ω is convex and

$$\forall x_1, x_2 \in \Omega,\ \ \lambda \in (0, 1),\ \ f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{2.3}$$

2.2 Basic Properties of Convex Functions

Definition 2.4. A function $f : \Omega \to \mathbb{R}$ is **concave** if and only if $-f$ is **convex**.

Proposition 2.1. Properties of convex functions:

(i) If f_1, f_2 are convex on Ω, so is $f_1 + f_2$;

(ii) If f is convex on Ω, then for any $a > 0$, af is also convex on Ω;

(iii) Any **sub-level/lower contour set** of a convex function f

$$\mathcal{L}(c) := \{x \in \mathbb{R}^n : f(x) \leq c\} \tag{2.4}$$

is convex.

Proof of (iii). Let $c \in \mathbb{R}$, and $x_1, x_2 \in SL(c)$. Let $s \in [0, 1]$. Since $x_1, x_2 \in \mathcal{L}(c)$, and $f(\cdot)$ is convex, $f(sx_1 + (1 - s)x_2) \leq sf(x_1) + (1 - s)f(x_2) \leq sc + (1 - s)c = c$. Which implies $sx_1 + (1 - s)x_2 \in \mathcal{L}(c)$. ∎

Example 2.1. ℓ_2 norm $f(x) : \mathbb{R}^n \to \mathbb{R} := ||x||_2$ is convex.

Proof. Note that for any $u, v \in \mathbb{R}^n$, by triangle inequality, $||u-(-v)|| \leq ||u-0||+||0-(-v)|| = ||u||+||v||$. Consequently, let $u, v \in \mathbb{R}^n$ and $s \in [0,1]$, then $||su+(1-s)v|| \leq ||su||+||(1-s)v|| = s||u||+(1-s)||v||$. Therefore, $||\cdot||$ is convex. ∎

Proposition 2.2. Any norm function $||\cdot||$ defined on a vector space $\mathcal{X}(\mathbb{R})$ is convex.

Proof. The proof follows the defining properties of norm,

$$\begin{aligned} ||\lambda x + (1-\lambda)y|| &\leq ||\lambda x|| + ||(1-\lambda)y|| && (2.5)\\ &= \lambda||x|| + (1-\lambda)||y|| && (2.6) \end{aligned}$$

∎

2.3 Characteristics of C^1 Convex Functions

Theorem 2.1 (C^1 criterions for convexity)**.** Let $f \in C^1$, then f is convex on a convex set Ω if and only if

$$\forall x, y \in \Omega,\ \ f(y) \geq f(x) + \nabla f(x) \cdot (y-x) \tag{2.7}$$

that is, *the linear approximation is never an overestimation of value of* f.

Proof. ($\Longrightarrow$) Suppose f is convex on a convex set Ω. Then $f(sy+(1-s)x) \leq sf(y)+(1-s)f(x)$ for every $x, y \in \Omega$ and $s \in [0,1]$, which implies, for every $s \in (0,1]$:

$$\frac{f(sy+(1-s)x)-f(x)}{s} \leq f(y)-f(x) \tag{2.8}$$

By taking the limit of $s \to 0$,

$$\begin{aligned} &\lim_{s\to 0} \frac{f(x+s(y-x))-f(x)}{s} \leq f(y)-f(x) && (2.9)\\ \Longrightarrow\ &\left.\frac{d}{ds}\right|_{s=0} f(x+s(y-x)) \leq f(y)-f(x) && (2.10)\\ &\Longrightarrow\ \nabla f(x)\cdot(y-x) \leq f(y)-f(x) && (2.11) \end{aligned}$$

($\Longleftarrow$) Let $x_0, x_1 \in \Omega$, let $s \in [0,1]$. Define $x^* := sx_0 + (1-s)x_1$, then

$$\begin{aligned} &f(x_0) \geq f(x^*) + \nabla f(x^*)\cdot(x_0 - x^*) && (2.12)\\ \Longrightarrow\ &f(x_0) \geq f(x^*) + \nabla f(x^*)\cdot[(1-s)(x_0-x_1)] && (2.13) \end{aligned}$$

Similarly,

$$\begin{aligned} &f(x_1) \geq f(x^*) + \nabla f(x^*)\cdot(x_1 - x^*) && (2.14)\\ \Longrightarrow\ &f(x_1) \geq f(x^*) + \nabla f(x^*)\cdot[s(x_1-x_0)] && (2.15) \end{aligned}$$

Therefore, $sf(x_0)+(1-s)f(x_1) \geq f(x^*)$. ∎

Theorem 2.2 (C^2 criterion for convexity)**.** $f \in C^2$ is a convex function on a convex set $\Omega \subseteq \mathbb{R}^n$ if and only if $\nabla^2 f(x) \succcurlyeq 0$ (i.e. positive semidefinite) for all $x \in \Omega$.

Corollary 2.1. When f is defined on $\mathbb{R}$, the C^2 criterion becomes $f''(x) \geq 0$.

Proof. ($\Longleftarrow$) Suppose $\nabla^2 f(x) \succcurlyeq 0$ for every $x \in \Omega$, let $x, y \in \Omega$. By the second order MVT,

$$f(y) = f(x) + \nabla f(x) \cdot (y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(x + s(y - x))(y - x) \text{ for some } s \in [0, 1] \tag{2.16}$$

$$\implies f(y) \geq f(x) + \nabla f(x) \cdot (y - x) \tag{2.17}$$

So f is convex by the C^1 criterion of convexity.
($\Longrightarrow$) Let $v \in \mathbb{R}^n$. Suppose, for contradiction, that for some $x \in \Omega$, $\nabla^2 f(x) \not\succcurlyeq 0$.
If such $x \in \partial\Omega$, note that $v^T \nabla^2 f(\cdot) v$ is continuous because $f \in C^2$, then there exists $\varepsilon > 0$ such that $\forall x' \in V_\varepsilon(x) \cap \Omega^{int},\ v^T \nabla^2 f(x') v < 0$.
Hence, one may assume with loss of generality that such $x \in \Omega^{int}$.
Because $x \in \Omega^{int}$, exists $\varepsilon' > 0$, such that $V_{\varepsilon'}(x) \subseteq \Omega^{int}$.
Define $\hat{v} := \frac{v}{\sqrt{\varepsilon'}}$, then for every $s \in [0, 1]$, $\hat{v}^T \nabla^2 f(x + s\hat{v})\hat{v} < 0$.
Let $y = x + \hat{v}$, by the mean value theorem,

$$f(y) = f(x) + \nabla f(x) \cdot (y - x) + \frac{1}{2}(y - x)^T \nabla^2 f[x + s(y - x)](y - x) \tag{2.18}$$

for some $s \in [0, 1]$.
This implies $f(y) < f(x) + \nabla f(x) \cdot (y - x)$, which contradicts the C^1 criterion for convexity. ∎

2.4 Minimum and Maximum of Convex Functions

Theorem 2.3. Let $\Omega \subseteq \mathbb{R}^n$ be a convex set, and $f : \Omega \to \mathbb{R}$ is a convex function. Let

$$\Gamma := \left\{ x \in \Omega : f(x) = \min_{x \in \Omega} f(x) \right\} \equiv \operatorname*{argmin}_{x \in \Omega} f(x) \tag{2.19}$$

If $\Gamma \neq \varnothing$, then

(i) Γ is convex;

(ii) any local minimum of f is the global minimum.

Proof (i). Let $x, y \in \Gamma$, $s \in [0, 1]$, then $sx + (1 - s)y \in \Omega$ because Ω is convex. Since f is convex, $f(sx + (1 - s)y) \leq sf(x) + (1 - s)f(y) = \min_{x \in \Omega} f(x)$. The inequality must be equality since it would contradicts the fact that $x, y \in \Gamma$. Therefore, $sx + (1 - s)y \in \Gamma$. ∎

Proof (ii). Let $x \in \Omega$ be a local minimizer for f, but assume, for contradiction, it is not a global minimizer. That is, there exists some other y such that $f(y) < f(x)$. Since f is convex,

$$f(x + t(y - x)) = f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y) < f(x) \tag{2.20}$$

for every $t \in (0, 1]$. Therefore, for every $\varepsilon > 0$, there exists $t^* \in (0, 1]$ such that $x + t^*(y - x) \in V_\varepsilon(x)$ and $f(x + t^*(y - x)) < f(x)$, this contradicts the fact that x is a local minimum. ∎

Theorem 2.4. Let $\Omega \subseteq \mathbb{R}^n$ be a <u>convex and compact</u> set, and $f : \Omega \to \mathbb{R}$ is a <u>convex</u> function. Then

$$\max_{x \in \Omega} f(x) = \max_{x \in \partial\Omega} f(x) \tag{2.21}$$

Proof. As we assumed, Ω is closed, therefore $\partial\Omega \subseteq \Omega$. Hence, $\max_{x \in \Omega} f \geq \max_{x \in \partial\Omega} f$.
Suppose, for contradiction, $\max_{x \in \Omega} f > \max_{x \in \partial\Omega} f$, then $x^* := \operatorname{argmax}_{x \in \Omega} f \in \Omega^{int}$.
Then we can construct a straight line through x^* and intersects $\partial\Omega$ at two points, $y_1, y_2 \in \partial\Omega$, such that

$x^* = sy_1 + (1-s)y_2$ for some $s \in (0,1)$. Further, since f is convex, $\max_{x\in\Omega} f(x) = f(x^*) \leq sf(y_1) + (1-s)f(y_2) \leq s\max_{\partial\Omega} f + (1-s)\max_{\partial\Omega} f = \max_{\partial\Omega} f$, which leads to a contradiction.
Therefore, $\max_{x\in\Omega} f = \max_{x\in\partial\Omega} f$. ■

Proposition 2.3. For $p, g > 1$ satisfying $\frac{1}{p} + \frac{1}{g} = 1$,

$$|ab| \leq \frac{1}{p}|a|^p + \frac{1}{g}|b|^g \tag{2.22}$$

Proof.

$$\begin{aligned} (-\log)|ab| &= (-\log)|a| + (-\log)|b| && (2.23)\\ &= \frac{1}{p}(-\log)|a|^p + \frac{1}{g}(-\log)|b|^p && (2.24)\\ (\because (-\log) \text{ is convex}) \quad &\geq (-\log)\left(\frac{1}{p}|a|^p + \frac{1}{g}|b|^p\right) && (2.25)\end{aligned}$$

And since $(-\log)$ is monotonically decreasing,

$$|ab| \leq \frac{1}{p}|a|^p + \frac{1}{g}|b|^p \tag{2.26}$$

■

Corollary 2.2.

$$|ab| \leq \frac{|a|^2 + |b|^2}{2} \tag{2.27}$$

3 Finite Dimensional Optimization

3.1 Unconstraint Optimization

Theorem 3.1 (Extreme Value Theorem)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and $K \subseteq \mathbb{R}^n$ be a compact set, then the minimization problem $\min_{x\in K} f(x)$ has a solution.

Remark 3.1. $f : \Omega \to \mathbb{R}$ is convex does not imply f is continuous.

Proposition 3.1. A convex function f defined on a convex open set is continuous.

Proof. Let $f : \Omega \to \mathbb{R}$ be a convex function, where $\Omega \subseteq \mathbb{R}^n$ is open. TODO: *Is this true?* ■

Proposition 3.2. A convex function f defined on an open interval in $\mathbb{R}$ is continuous.

Proof. See homework 1. The proof involves squeeze theorem. ■

Proof of EVT.. Let $f : K \to \mathbb{R}$ be a continuous function defined on a compact set K.
WLOG, we only prove the existence of $\min f$, since the existence of max can be easily proven by applying the exact same argument on $-f$.
That is, we claim the infimum of $f(K)$ is attained within K.
Because K is compact, the continuity of f implies $f(K)$ is compact.
By the completeness axiom of $\mathbb{R}$, $m := \inf_{x\in K} f(x)$ is well-defined. There exists a sequence $(x_i) \subseteq K$, such that $(f(x_i)) \to m$. Because K is compact, there exists a subsequence (x_{ik}) of (x_i) converges to some limit $x^* \in K$.

Since f is continuous, $(f(x_{ik})) \to f(x^*)$, which is a subsequence of the convergent sequence $(f(x_i))$, and they must converge to the same limit. Hence, $f(x^*) = m$, and the infimum is attained at $x^* \in K$. ■

Theorem 3.2 (Heine–Borel)**.** Let $K \subseteq \mathbb{R}^n$, then the following are equivalent:

(i) K is compact (every open cover of K has a finite sub-cover);

(ii) K is closed and bounded;

(iii) Every sequence in K has a convergent subsequence converges to a point in K.

Proposition 3.3. Let $\{h_i\}$ and $\{g_i\}$ be sets of continuous functions on $\mathbb{R}^n$, the the set of all points in $\mathbb{R}^n$ that satisfy

$$\begin{cases} h_i(x) = 0 \ \forall i \\ g_j(x) \leq 0 \ \forall j \end{cases} \tag{3.1}$$

is a closed set. Moreover, if the qualified set is also bounded, then it is compact.

Proof. For every equality constraint h_i, it can be represented as the conjunction of two inequality constraint, namely $h_i^\alpha(x) := -h_i(x) \leq 0 \wedge h_i^\beta(x) := h_i(x) \leq 0$. Then the constraint collection is equivalent to

$$\begin{cases} h_i^\alpha(x) \leq 0 \ \forall i \\ h_i^\beta(x) \leq 0 \ \forall i \\ g_j(x) \leq 0 \ \forall j \end{cases} \tag{3.2}$$

The subset of $\mathbb{R}^n$ qualified by each individual constraint is closed by the property of continuous functions (i.e. a continuous function's pre-image of closed set is closed). And the intersection of arbitrarily many closed sets is closed. ■

Remark 3.2. Computer algorithms for solving minimization problems try to construct a sequence of (x_i) such that $f(x_i)$ decreases to $\min f$ rapidly.

The optimization problems investigated in this section can be formulated as

$$\min_{x \in \Omega} f(x) \tag{3.3}$$

where $\Omega \subseteq \mathbb{R}^n$. Typically, for simplicity, Ω are often $\mathbb{R}^n$, an open subset of R^n, or the closure of some open subset of $\mathbb{R}^n$.

Everything above minimization discussed in this section is applicable to maximization as well using the proposition below.

Proposition 3.4. When $\Omega = \mathbb{R}^n$, the unconstrained minimization has the following properties

(i) $\operatorname{argmax} f = \operatorname{argmin}(-f)$;

(ii) $\max f = -\min(-f)$

Proof. Immediate by applying definitions of maximum and minimum. ■

Definition 3.1. A function $f : \Omega \to \mathbb{R}$ has **local minimum** at $x_0 \in \Omega$ if

$$\exists \varepsilon > 0 \ s.t. \ \forall x \in V_\varepsilon(x_0) \cap \Omega, \ f(x_0) \leq f(x) \tag{3.4}$$

f attains **strictly local minimum** at x_0 if

$$\exists \varepsilon > 0 \ s.t. \ \forall x \in V_\varepsilon(x_0) \cap \Omega \backslash \{x_0\} \ f(x_0) < f(x) \tag{3.5}$$

f attains **global minimum** at x_0 if

$$\forall x \in \Omega \ f(x_0) \leq f(x) \tag{3.6}$$

f attains **strict global minimum** at x_0 if

$$\forall x \in \Omega \backslash \{x_0\} \ f(x_0) < f(x) \tag{3.7}$$

Note that strict global minimum is always unique.

Theorem 3.3 (Necessary Condition for Local Minimum)**.** Let $C^1 \ni f : \Omega \to \mathbb{R}$, let $x_0 \in \Omega$ be a local minimum of f, then for every *feasible direction* v at x_0,

$$\nabla f(x_0) \cdot v \geq 0 \tag{3.8}$$

This theorem serves as the primary defining property of local minimum.

Definition 3.2. For $x_0 \in \Omega \subseteq \mathbb{R}^n$, $v \in \mathbb{R}^n$ is a **feasible direction** at x_0 if

$$\exists \overline{s} > 0 \ s.t. \ \forall s \in [0, \overline{s}], x_0 + sv \in \Omega \tag{3.9}$$

Proof of Necessary Condition. Let $x_0 \in \Omega$ be a local minimum, and let v be a Define auxiliary function $g(s) := f(x + sv)$. And since g attains minimum at $s = 0$, there exists some $\overline{s} > 0$ such that

$$g(s) - g(0) \geq 0 \ \forall s \in [0, \overline{s}] \tag{3.10}$$

Therefore

$$g'(0) := \lim_{s \to 0} \frac{g(s) - g(0)}{s - 0} \geq 0 \tag{3.11}$$

The alternative form of derivative can be derived using chain rule as

$$g'(0) = \nabla f(x + sv) \cdot v \mid_{s=0} = \nabla f(x) \cdot v \tag{3.12}$$

By combing the two identities above, $\nabla f(x) \cdot v \geq 0$. ■

Alternative Proof of Necessary Condition (not that rigorous). The prove is almost immediate, if there exists a feasible direction v^* such that $\nabla f(x_0) \cdot v^* < 0$, for every $\varepsilon > 0$, one can construct $x' := x^* + sv^*$ with sufficiently small s so that $x' \in V_\varepsilon(x^*) \cap \Omega$ and $f(x') < f(x^*)$. ■

Corollary 3.1. When Ω is open, then x_0 is a local minimum $\implies \nabla f(x_0) = 0$.

Proof. Since Ω is open, any sufficiently small $v \neq 0$ such that both v and $-v$ are feasible directions at x_0, applying the necessary condition on both v and $-v$ provides the equality. ■

Theorem 3.4 (Second Order Necessary Condition for Local Minimum)**.** Let $C^2 \ni f : \Omega \to \mathbb{R}$, let $x_0 \in \Omega$ be a local minimum of f, then for every non-zero feasible direction v at x_0,

(i) $\nabla f(x_0) \cdot v \geq 0$;

(ii) $\nabla f(x_0) \cdot v = 0 \implies v^T \nabla^2 f(x_0) v \geq 0$.

Proof. Let x_0 be a local minimum and v be a feasible direction at Ω, and $s \in (0, \bar{s}]$. The first statement is the immediate result of the first order necessary condition. Now suppose $\nabla f(x_0) = 0$, by the Taylor's theorem,

$$0 \leq f(x_0 + sv) - f(x_0) = s\nabla f(x_0) \cdot v + \frac{s^2}{2} v^T \nabla^2 f(x_0) v + o(s^2) \tag{3.13}$$

$$= \frac{s^2}{2} v^T \nabla^2 f(x_0) v + o(s^2) \tag{3.14}$$

Since $s^2 > 0$, divide both sides by s^2 and take limit,

$$\lim_{s \to 0} \frac{f(x_0 + sv) - f(x_0)}{s^2} = \lim_{s \to 0} \left\{ \frac{1}{2} v^T \nabla^2 f(x_0) v + \frac{o(s^2)}{s^2} \right\} \tag{3.15}$$

$$= \frac{1}{2} v^T \nabla^2 f(x_0) v + \lim_{s \to 0} \frac{o(s^2)}{s^2} \tag{3.16}$$

$$= \frac{1}{2} v^T \nabla^2 f(x_0) v \geq 0 \tag{3.17}$$

■

Example 3.1. $f(x, y) = x^2 - xy + y^2 - 3y : \Omega = \mathbb{R}^2 \to \mathbb{R}$. Then at $(x_0, y_0) = (1, 2)$,

$$\nabla f(x_0, y_0) = (2x_0 - y, -x_0 + 2y_0 - 3) = (0, 0) \tag{3.18}$$

$$\nabla^2 f(x_0, y_0) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \succcurlyeq 0 \tag{3.19}$$

Definition 3.3. Let $A \in \mathbb{R}^{n \times n}$, A is

(i) **Positive definite** (denoted as $A \succ 0$) if $x^T A x > 0 \ \forall x \neq 0$, if and only if all eigenvalues $\lambda_i > 0$;

(ii) **Positive Semi-definite** (denoted as $A \succcurlyeq 0$) if $x^T A x \geq \ \forall x \in \mathbb{R}^n$, if and only if all eigenvalues $\lambda_i \geq 0$.

Theorem 3.5 (Sylvester's Criterion)**.** Let $A \in \mathbb{R}^{n \times n}$ be a Hermitian matrix (i.e. $A = \overline{A^T}$)[1], then

1. $A \succ 0 \iff$ all *leading principal minors* have positive determinants;
2. $A \succcurlyeq 0 \iff$ all leading principal minors have non-negative determinants.

Theorem 3.6 (Second Order Sufficient Condition for Interior Local Minima)**.** Let $f : C^2(\Omega, \mathbb{R})$, for some $x_0 \in \Omega$, if

(i) $\nabla f(x_0) = 0$,

(ii) *(and)* $\nabla^2 f(x_0) \succ 0$.

then x_0 is a <u>strictly local minimizer</u>.

Lemma 3.1. Suppose $\nabla^2 f(x_0)$ is positive definite, then

$$\exists \ a > 0 \ s.t. \ v^T \nabla^2 f(x_0) v \geq a||v||^2 \quad \forall v \tag{3.20}$$

That is, the quadratic form of a positive definite matrix is bounded away from zero.

[1] $\overline{A^T}$ denotes the complex conjugate of the transpose, a matrix with *real entries* is Hermitian if and only if it is symmetric.

Proof of the Lemma. Recall that a squared matrix Q is called **orthogonal** when every column and row of it is an orthogonal unit vector. So that for every orthogonal matrix Q, $Q^TQ = I$, which implies $Q^T = Q^{-1}$. Further, note that

$$||Qv||^2 = (Qv)^T(Qv) = v^TQ^TQv = ||v||^2 \tag{3.21}$$

$$\implies ||Qv|| = ||v|| \ \forall v \in \mathbb{R}^n \tag{3.22}$$

Let $v \in \mathbb{R}^n$, consider the eigenvector decomposition of $\nabla^2 f(x_0)$, let w satisfy $v = Qw$:

$$Q^T\nabla^2 f(x_0)Q = \text{diag}(\lambda_1, \cdots, \lambda_n) \tag{3.23}$$

$$\implies v^T\nabla^2 f(x_0)v = (Qw)^T\nabla^2 f(x_0)(Qw) \tag{3.24}$$

$$= w^TQ^T\nabla^2 f(x_0)Qw \tag{3.25}$$

$$= w^T\text{diag}(\lambda_1, \cdots, \lambda_n)w \tag{3.26}$$

$$= \lambda_1 w_1^2 + \cdots + \lambda_n w_n^2 \tag{3.27}$$

Let $a := \min\{\lambda_1, \cdots, \lambda_n\}$,

$$... \geq a||w||^2 = a||Q^Tv||^2 = a||v||^2 \tag{3.28}$$

■

Proof of the Theorem. Let $x \in \Omega$, suppose $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succneqq 0$. By the second order Taylor approximation,

$$f(x_0 + v) - f(x_0) = \nabla f(x_0)^Tv + \frac{1}{2}v^T\nabla^2 f(x_0)v + o(||v||^2) \tag{3.29}$$

$$= \frac{1}{2}v^T\nabla^2 f(x_0)v + o(||v||^2) \tag{3.30}$$

$$\geq \frac{a}{2}||v||^2 + o(||v||^2) \text{ for some } a > 0 \tag{3.31}$$

$$= ||v||^2\left(\frac{a}{2} + \frac{o(||v||^2)}{||v||}\right) \tag{3.32}$$

$$> 0 \ \text{ for sufficiently small } v \tag{3.33}$$

Therefore, $f(x_0) < f(x) \ \forall x \in V_\varepsilon(x_0)$. ■

3.2 Equality Constraints: Lagrangian Multiplier

3.2.1 Tangent Space to a (Hyper) Surface at a Point

Definition 3.4. A **surface** $\mathcal{M} \subseteq \mathbb{R}^n$ is defined as

$$\mathcal{M} := \{x \in \mathbb{R}^n : h_i(x) = 0 \ \forall i\} \tag{3.34}$$

where h_i are all C^1 functions.

Definition 3.5. A **differentiable curve** on a surface $\mathcal{M}$ is a C^1 function mapping from $(-\varepsilon, \varepsilon)$ to $\mathcal{M}$. *Remark: in previous calculus courses, differentiable curves are often referred to as parameterizations.*

Let $x(s)$ be a differentiable curve on $\mathcal{M}$ passes through $x_0 \in \mathcal{M}$, re-parameterize it so that $x(0) = x_0$.

Then vector

$$v := \left.\frac{d}{ds}\right|_{s=0} x(s) \tag{3.35}$$

touches $\mathcal{M}$ *tangentially.*

Definition 3.6. Any vector v generated by some differentiable curve on $\mathcal{M}$ and takes above form is a **tangent vector** on $\mathcal{M}$ through x_0.

Definition 3.7. The **tangent space** to $\mathcal{M}$ at x_0 is defined to be the set of all tangent vectors:

$$T_{x_0}\mathcal{M} := \left\{ v \in \mathbb{R}^n : v := \left.\frac{d}{ds}\right|_{s=0} x(s) \text{ for some } x \in C^1((-\varepsilon, \varepsilon), \mathcal{M}) \ s.t.\ x(0) = x_0 \right\} \tag{3.36}$$

Example 3.2. Define

$$\mathcal{M} := \left\{ x \in \mathbb{R}^2 : ||x||_2 = 1 \right\} \tag{3.37}$$

By defining C^1 functions $g(x) := ||x||_2^2 - 1$, $\mathcal{M}$ is a surface. The tangent space of $\mathcal{M}$ at x_0 is

$$T_{x_0}\mathcal{M} = \{ v \in \mathbb{R}^n : \langle v, x_0 \rangle = 0 \} \tag{3.38}$$

Definition 3.8. Let $\mathcal{M}$ be a surface defined using C^1 functions, a point $x_0 \in \mathcal{M}$ is a **regular point** of the constraints if

$$\{ \nabla h_1(x_0), \cdots, \nabla h_k(x_0) \} \tag{3.39}$$

are linearly independent.
Remark: if there is only one constraint h, then x_0 is regular if and only if $\nabla h(x_0) \neq 0$.

Notation 3.1. Define the T space on equality constraint as

$$T_{x_0} := \{ x \in \mathbb{R}^n : \langle x_0, \nabla h_i(x_0) \rangle = 0 \ \forall i \in [k] \} \tag{3.40}$$

Example 3.3 (Counter example)**.** Define

$$\mathcal{M} := \left\{ (x, y) \in \mathbb{R}^2 : h(x, y) = xy = 0 \right\} \tag{3.41}$$

Then it is easy to verify that $(0, 0)$ is not a regular point. And

$$T_{0,0} = \{ (x, y) \in \mathbb{R}^2 : (x, y) \cdot (0, 0) = 0 \} = \mathbb{R}^2 \tag{3.42}$$

$$\neq T_{0,0}\mathcal{M} = \{ (x, y) \in \mathbb{R}^2 : x = 0 \vee y = 0 \} \tag{3.43}$$

Theorem 3.7. Suppose x_0 is a *regular point* of $\mathcal{M} := \{ h_i(x) = 0, i = 1, \cdots, k \}$, then $T_{x_0} = T_{x_0}\mathcal{M}$.

Proof. Show $T_{x_0}\mathcal{M} \subseteq T_{x_0}$.
Suppose x_0 is a regular point of $\mathcal{M}$. Let $v \in T_{x_0}\mathcal{M}$, then there exists some differentiable curve $x(\cdot) : V_\varepsilon(0) \to \mathcal{M}$ such that $x(0) = x_0$, such that

$$v = \left.\frac{d}{ds}\right|_{s=0} x(s) \tag{3.44}$$

Note that $h_i(x(s)) = 0$ is constant for every $i \in [k]$, therefore

$$\left.\frac{d}{ds}\right|_{s=0} h_i(x(s)) \tag{3.45}$$

By the chain rule,

$$\nabla h_i(x_0) \cdot v = 0 \ \forall i \tag{3.46}$$

Therefore $v \in T_{x_0}$.
Show $T_{x_0} \subseteq T_{x_0}\mathcal{M}$.

(i) x_0 is regular $\implies T_{x_0}\mathcal{M}$ is a vector space;

(ii) $T_{x_0} = \text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\}^\perp$.

Show $T_{x_0} \subseteq \text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\}^\perp$:
Let $v \in T_{x_0}$, then $v \perp \nabla h_i(x_0)$ for every i. Therefore v is orthogonal to every linear combination of $\nabla h_i(x_0)$, and therefore orthogonal to the span.
Show $\text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\}^\perp \subseteq T_{x_0}$:
Let v in the perp of the span, then v is orthogonal to every basis of the span, so $v \in T_{x_0}$. ∎

Lemma 3.2. Let $f, h_1, \cdots, h_k \in C^1$ defined on open subset $\Omega \subseteq \mathbb{R}^n$. Define $\mathcal{M} := \{x \in \mathbb{R}^n : h_i(x) = 0 \ \forall i\}$. Suppose $x_0 \in \mathcal{M}$ is a local minimum of f on $\mathcal{M}$, then

$$\nabla f(x_0) \perp T_{x_0}\mathcal{M} \tag{3.47}$$

Proof. WLOG $\Omega = \mathbb{R}^n$, take $v \in T_{x_0}\mathcal{M}$. Then there exists some differentiable curve x on $\mathcal{M}$ satisfying $v = x'(0)$. Because x_0 is a local minimum of f on Ω, $s = 0$ is a local minimum of $f(x(s))$, moreover, it is an interior minimum. By chain rule and the necessary condition of local minimum,

$$Df(x(0)) = \nabla f(x(0)) \cdot x'(0) = 0 \tag{3.48}$$
$$\implies \nabla f(x_0) \cdot v = 0 \tag{3.49}$$

Therefore $\nabla f(x_0) \perp T_{x_0}\mathcal{M}$. ∎

Theorem 3.8 (Lagrange Multipliers: First Order Necessary Condition)**.** Let $f, h_1, \cdots, h_k \in C^1$ defined on open subset $\Omega \subseteq \mathbb{R}^n$. Let x_0 be a regular point of the constraint set $\mathcal{M} := \bigcap_{i=1}^k h_i^{-1}(0)$. Suppose x_0 is a local minimum of $\mathcal{M}$, then there exists $\lambda_1, \cdots, \lambda_k \in \mathbb{R}$ such that

$$\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) = 0 \tag{3.50}$$

Remark: if we define Lagrangian $\mathcal{L}(x, \lambda_i) := f(x) + \sum_{i=1}^k h_i(x)$, then the theorem says the local minimum is a critical point of $\mathcal{L}$.

Proof. Because x_0 is a regular point, then by previous lemma, $\nabla f(x_0) \perp T_{x_0}\mathcal{M}$. Moreover,

$$T_{x_0}\mathcal{M} = T_{x_0} = (\text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\})^\perp \tag{3.51}$$

Also, because x_0 is a local minimum,

$$\nabla f(x_0) \perp T_{x_0}\mathcal{M} \tag{3.52}$$

Therefore, $\nabla f(x_0) \in (T_{x_0}\mathcal{M})^{\perp} = (\text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\})^{\perp\perp} = \text{span}\{\nabla h_1(x_0), \cdots, \nabla h_k(x_0)\}$, where the last equality holds in finite dimensional cases. Hence, it is obvious that we can write $\nabla f(x_0)$ as a linear combination of $\{\nabla h_i(x_0)\}$. ■

Theorem 3.9 (Second Order Necessary Condition)**.** Let $f, h_i \in C^2$, if x_0 is a local minimum on previously defined surface $\mathcal{M}$, then there exists Lagrangian multipliers $\{\lambda_i\}$ such that

(i) $\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) = 0$ ($\nabla_x \mathcal{L} = 0$);

(ii) And $\nabla^2 f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla^2 h_i(x_0) \succcurlyeq 0$ on $T_{x_0}\mathcal{M}$ ($\nabla_x^2 \mathcal{L} \succcurlyeq 0$).

Remark: whenever x_0 is a local minimum, it must be a critical point of $\mathcal{L}$, and $\mathcal{L}$ is positive semidefinite on the tangent space at x_0.

Proof. The first result is exactly the same as the first order condition proven above.
To show the second result, let $x(s) \in \mathcal{M}$ be an arbitrary differentiable curve on $\mathcal{M}$ such that $x(0) = x_0$. Then,

$$\frac{d}{ds} f(x(s)) = \nabla f(x(s)) \cdot x'(s) \tag{3.53}$$
$$\frac{d^2}{ds^2} f(x(s)) = x'(s)^T \nabla^2 f(x(s)) x'(s) + \nabla f(x(s)) x''(s) \tag{3.54}$$

By the second order Taylor theorem, for every s such that $x(s) \in \mathcal{M}$,

$$f(x(s)) - f(x_0) = s \nabla f(x_0) \cdot x'(0) + \frac{s^2}{2}\left[x'(0)^T \nabla^2 f(x(0)) x'(s) + \nabla f(x(0)) x''(0)\right] + o(s^2) \tag{3.55}$$

Note that by definition, $x'(0)$ is in the tangent space at x_0. Also, we've shown previously that $\nabla f(x_0)$ is orthogonal to the tangent space at x_0, therefore,

$$f(x(s)) - f(x_0) = \frac{s^2}{2}\left[x'(0)^T \nabla^2 f(x(0)) x'(s) + \nabla f(x(0)) x''(0)\right] + o(s^2) \tag{3.56}$$

Also, by the definition of $\mathcal{M}$, all constraints hold with equality:

$$f(x_0) = f(x_0) + \sum_{i=1}^{k} \lambda_i h_i(x_0) \tag{3.57}$$

where λ_i's are from the first result. Hence,

$$f(x(s)) - f(x_0) = \frac{s^2}{2}\left[x'(0)^T \left(\nabla^2 f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla^2 h_i(x_0)\right) x'(0) + \left(\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0)\right) x''(0)\right] + o(s^2) \tag{3.58}$$

$$= \frac{s^2}{2} x'(0)^T \left(\nabla^2 f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla^2 h_i(x_0)\right) x'(0) + o(s^2) \tag{3.59}$$

And above expression is greater or equal to zero because x_0 is a local minimum,

$$\frac{s^2}{2}x'(0)^T\left(\nabla^2 f(x_0)+\sum_{i=1}^{k}\lambda_i\nabla^2 h_i(x_0)\right)x'(0)+o(s^2)\geq 0 \tag{3.60}$$

$$\implies x'(0)^T\left(\nabla^2 f(x_0)+\sum_{i=1}^{k}\lambda_i\nabla^2 h_i(x_0)\right)x'(0)+\frac{o(s^2)}{s^2}\geq 0 \tag{3.61}$$

$$\overset{s\to 0}{\implies} x'(0)^T\left(\nabla^2 f(x_0)+\sum_{i=1}^{k}\lambda_i\nabla^2 h_i(x_0)\right)x'(0)\geq 0 \tag{3.62}$$

Where $x'(0)$ is a vector in the tangent space at x_0 by definition. Moreover, the curve $x(s)$ was chosen arbitrarily, so the argument works for every curve and therefore every tangent vector, and what's desired is shown. ■

Example 3.4.

$$\min f(x,y)=x^2-y^2 \tag{3.63}$$

$$s.t.\ h(x,y)=y=0 \tag{3.64}$$

First order condition suggests $(x_0,y_0)=(0,0)$ Note that the tangent space at (x_0,y_0) is $\text{span}\{\nabla h_i\}^{\perp}$:

$$T_{x_0}\mathcal{M}=\{(u,0):u\in\mathbb{R}\} \tag{3.65}$$

and

$$\nabla_x^2\mathcal{L}=\begin{pmatrix}2 & 0\\ 0 & -2\end{pmatrix} \tag{3.66}$$

is obviously positive semidefinite (actually positive definition) on the tangent space.

Theorem 3.10 (Second Order Sufficient Conditions)**.** Let $f,h_i\in C^2$ on open $\Omega\subseteq\mathbb{R}^n$, and $x_0\in\mathcal{M}$ is a regular point, if there exists $\lambda_i\in\mathbb{R}$ such that

(i) $\nabla_x\mathcal{L}(x_0,\lambda_i)=0$;

(ii) $\nabla_x^2\mathcal{L}(x_0,\lambda_i)\succ 0$ on $T_{x_0}\mathcal{M}$,

then x_0 is a *strict* local minimum.

Proof. Recall that $\nabla^2 f(x_0)+\sum\lambda_i\nabla^2 h_i(x_0)$ positive definite on $T_{x_0}\mathcal{M}$ implies there exists $a>0$ (a is taken to be equal to the least eigenvalue of $\nabla_x^2\mathcal{L}$) such that

$$v^T[\nabla^2 f(x_0)+\sum\lambda_i\nabla^2 h_i(x_0)]v\geq a||v||^2\quad\forall v\in T_{x_0}\mathcal{M} \tag{3.67}$$

Let $x(s)\in\mathcal{M}$ be a curve such that $x(0)=x_0$ and $v=x'(0)$. WLOG, $||x'(0)||=1$. By the second order

Taylor expansion,

$$
\begin{aligned}
f(x(s)) - f(x(0)) &= s \left.\frac{d}{ds}\right|_{s=0} f(x(s)) + \frac{s^2}{2} \left.\frac{d^2}{ds^2}\right|_{s=0} f(x(s)) + o(s^2) && (3.68)\\
&= s \left.\frac{d}{ds}\right|_{s=0} \left[f(x(s)) + \sum \lambda_i h_i(x(s))\right] + \frac{s^2}{2} \left.\frac{d^2}{ds^2}\right|_{s=0} \left[f(x(s)) + \sum \lambda_i h_i(x(s))\right] + o(s^2) && (3.69)\\
&= s\nabla_x \mathcal{L}(x_0, \lambda_i) \cdot x'(0) + \frac{s^2}{2} \left[x'(0)^T \nabla_x^2 \mathcal{L}(x_0, \lambda_i) x'(0) + \nabla_x \mathcal{L}(x_0, \lambda_i) x''(0)\right] + o(s^2) && (3.70)\\
&= \frac{s^2}{2} x'(0)^T \nabla_x^2 \mathcal{L}(x_0, \lambda_i) x'(0) + o(s^2) && (3.71)\\
&\geq \frac{s^2}{2} a||x'(0)||^2 + o(s^2) \quad \text{where } a > 0 && (3.72)\\
&= s^2 \left(\frac{a}{2} + \frac{o(s^2)}{s^2}\right) && (3.73)\\
&\overset{s\to 0}{>} 0 && (3.74)
\end{aligned}
$$

Therefore, for sufficiently small s, $f(x(s)) - f(x(0)) > 0$. And this is true for every curve x on $\mathcal{M}$. So $x(0)$ is a strict local minimum. ■

3.3 Remark on the Connection Between Constrained and Unconstrained Optimizations

Example 3.5. Consider

$$
\begin{aligned}
&\min f(x, y, z) && (3.75)\\
&s.t. g(x, y, z) = z - h(x, y) = 0 && (3.76)
\end{aligned}
$$

where $\mathcal{M}$ is the graph of h. Using Lagrangian multiplier provides necessary condition: $\nabla f + \lambda \nabla g = 0$,

$$
\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} + \lambda \begin{pmatrix} -h_x \\ -h_y \\ 1 \end{pmatrix} = 0 \qquad (3.77)
$$

Convert the constrained optimization into an unconstrained optimization as

$$
\min_{(x,y)\in\mathbb{R}^2} F(x, y) = f(x, y, h(x, y)) \qquad (3.78)
$$

The necessary condition for unconstrained optimization is

$$
\begin{aligned}
\nabla F(x, y) &= \begin{pmatrix} f_x + f_z h_x \\ f_y + f_z h_y \end{pmatrix} && (3.79)\\
&= \begin{pmatrix} f_x \\ f_y \end{pmatrix} - f_z \begin{pmatrix} -h_x \\ -h_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} && (3.80)
\end{aligned}
$$

Define $\lambda := -f_z$.

$$\nabla F(x,y) = \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} + \lambda \begin{pmatrix} -h_x \\ -h_y \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3.81}$$

3.4 Inequality Constraints

Definition 3.9. Let x_0 satisfy the set of constraints

$$(\dagger) \begin{cases} h_i(x) & = 0 \quad i \in \{1, \cdots, k\} \\ g_j(x) & \leq 0 \quad j \in \{1, \cdots, \ell\} \end{cases} \tag{3.82}$$

we say that the constraint g_i is **active** at x_0 if $g_i(x_0) = 0$, and is **inactive** at x_0 if $g_i(x_0) < 0$.

Definition 3.10. Split the collection of inequality constraints into active and inactive constraints, let $\Theta(x_0)$ denote the collection of active indices, that's:

$$g_j(x_0) = 0 \ \forall j \in \Theta(x_0) \tag{3.83}$$

$$g_j(x_0) < 0 \ \forall j \notin \Theta(x_0) \tag{3.84}$$

Then x_0 is said to be a **regular point** of the constraint if

$$\{\nabla h_i(x_0) \ \forall i \in \{1, \cdots, k\}; \underbrace{\nabla g_j(x_0) \ \forall j \in \Theta(x_0)}_{\text{Active Constraints}}\} \tag{3.85}$$

is linearly independent.

Theorem 3.11 (The First Order Necessary Condition for Local Minimum: Kuhn-Tucker Conditions)**.** Let Ω be an open subset of $\mathbb{R}^n$ with constraints h_i and g_i to be C^1 on Ω. Suppose $x_0 \in \Omega$ is a regular point with respect to constraints, further suppose x_0 is a local minimum, then there exists some $\lambda_i \in \mathbb{R}$ and $\mu_j \in \mathbb{R}_+$ such that

(i) $\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) + \sum_{j=1}^\ell \mu_j \nabla g_j(x_0) = 0$ (i.e. $\nabla_x \mathcal{L}(x, \lambda, \mu) = 0$);

(ii) $\mu_j g_j(x_0) = 0$ (*Complementary slackness*).

Remark 1: by complementary slackness, all μ_j corresponding to inactive inequality constraints are zero.
Remark 2: it is possible for an active constraint to have zero multiplier.

Proof. Let x_0 be a local minimum for f satisfying constraints, equivalently, it is a local minimum for equality constraints and active inequality constraints.
By the first order necessary condition for local minimum with equality constraints, there exists $\lambda_i, \mu_j \in \mathbb{R}$ such that

$$\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) + \sum_{j \in \Theta(x_0)} \mu_j \nabla g_j(x_0) = 0 \tag{3.86}$$

Then by setting $\mu_j = 0$ for all $j \notin \Theta(x_0)$ one have

$$\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) + \sum_{j=1}^\ell \mu_j \nabla g_j(x_0) = 0 \tag{3.87}$$

By construction, the complementary slackness is satisfied. At this stage, we have construct $\lambda_i \in \mathbb{R}$ and $\mu_j \in \mathbb{R}$ satisfying both conditions, we still need to argue that $\mu_j \geq 0$ for every j. ■

Theorem 3.12 (The Second Order Necessary Conditions)**.** Let Ω be an open subset of $\mathbb{R}^n$, and $f, h_1, \cdots, h_k, g_1, \cdots, g_\ell \in C^2(\mathbb{R}^n, \mathbb{R})$. Let x_0 be a regular point of the constraints (†). Suppose x_0 is a local minimum of f subject to constraint (†), then there exists $\lambda_i \in \mathbb{R}$ and $\mu_j \geq 0$ such that

(i) $\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) + \sum_{j=1}^\ell \mu_j \nabla g_j(x_0) = 0$;

(ii) $\mu_j g_j(x_0) = 0$;

(iii) $\nabla^2 f(x_0) + \sum_{i=1}^k \lambda_i \nabla^2 h_i(x_0) + \sum_{j=1}^\ell \mu_j \nabla^2 g_j(x_0)$ is positive semidefinite on the tangent space to activate constraints at x_0.

Proof. (i) and (ii) are immediate result from the first order necessary condition.
Suppose x_0 is a local minimum for (†), then x_0 is a local minimum for active constraints at x_0.
Therefore, $\nabla^2 \hat{\mathcal{L}} = \nabla^2 f(x_0) + \sum_{i=1}^k \lambda_i \nabla^2 h_i(x_0) + \sum_{j \in I(x_0)} \mu_j \nabla^2 g_j(x_0)$ is positive semidefinite on the tangent space to active constraints. Note that because $\mu_j = 0$ for inactive constraints, therefore $\nabla^2 \hat{\mathcal{L}} = \nabla^2 \mathcal{L}$ at x_0, and both of them are positive semidefinite on the tangent space corresponding to active constraints. ■

Theorem 3.13 (The Second Order Sufficient Conditions)**.** Let Ω be an open subset of $\mathbb{R}^n$, let $f, h_i, q_j \in C^2(\Omega)$. Consider minimizing $f(x)$ with the constraint

$$(\dagger) \begin{cases} h_i(x) = 0 & \forall i \\ g_j(x) \leq 0 & \forall j \\ x \in \Omega \end{cases} \tag{3.88}$$

Suppose there exists a feasible x_0 satisfying (†) and $\lambda_i \in \mathbb{R}$ and $\mu_j \in \mathbb{R}_+$ such that

(i) $\nabla f(x_0) + \sum_{i=1}^k \lambda_i \nabla h_i(x_0) + \sum_{j=1}^\ell \mu_j \nabla g_j(x_0) = 0$;

(ii) $\mu_j g_j(x_0) = 0$ (*Complementary slackness*).

If the Hessian matrix for Lagrangian $\nabla_x^2 \mathcal{L}(x_0)$ is positive definite on $\tilde{\tilde{T}}_{x_0}$, the space of **strongly active** constraints at x_0, then x_0 is a strict local minimum.

Definition 3.11. A constraint g_j is **strongly active** at x_0 if $g_j(x_0) = 0$ (so it is active) and $\mu_j > 0$.

Notation 3.2. For convenience, we can rearrange the collection of constraints such that, among the ℓ constrains in total, the first ℓ' constraints are *active* at x_0 and the first ℓ'' constraints are *strongly active.* Note that $\ell'' \leq \ell' \leq \ell$.
Define

$$\tilde{T}_{x_0} := \{v \cdot \nabla h_i(x_0) = 0\ \forall i \wedge v \cdot \nabla g_j(x_0) \text{ for all } g_j \text{ active.}\} \tag{3.89}$$

$$\tilde{\tilde{T}}_{x_0} := \{v \cdot \nabla h_i(x_0) = 0\ \forall i \wedge v \cdot \nabla g_j(x_0) \text{ for all } g_j \text{ strongly active.}\} \tag{3.90}$$

Clearly, $\tilde{T}_{x_0} \subseteq \tilde{\tilde{T}}_{x_0}$ because there are (weakly) more active constraints than strongly active constraints.

Proof of the Sufficient Condition. Suppose, for contradiction, x_0 is not a strict local minimum.
Claim 1: There exists unit vector $v \in \mathbb{R}^n$ such that

(a) $\nabla f(x_0) \cdot v \leq 0$;

(b) $\nabla h_i(x_0) \cdot v = 0$ for every i;

(c) $\nabla g_j(x_0) \cdot v \leq 0$ for all $j \leq \ell'$ (active constraints).

Proof of Claim 1. Because x_0 is not a strictly local minimum, one can construct a sequence of feasible points $(x_k) \to x_0$ by setting $\varepsilon = \frac{1}{k}$ for every $k \in \mathbb{N}$ such that $f(x_k) \leq f(x_0)$.
Let $v_k := \frac{x_k - x_0}{||x_k - x_0||}$, $s_k := ||x_k - x_0||$. Note that every v_k is in unit sphere, which is compact. Therefore, there exists a subsequence of (v_k) converges to some unit vector v.

$$0 \geq f(x_k) - f(x_0) = f(x_0 + s_k v_k) - f(x_0) \ \forall k \in \mathbb{N} \tag{3.91}$$

The first order Taylor series suggests the following holds for every $k \in \mathbb{N}$:

$$0 \geq f(x_0 + s_k v_k) - f(x_0) \tag{3.92}$$
$$= s_k \nabla f(x_0) \cdot v_k + o(s_k) \tag{3.93}$$
$$0 = h_i(x_0 + s_k v_k) - h_i(x_0) = s_k \nabla h_i(x_0) \cdot v_k + o(s_k) \tag{3.94}$$
$$0 \geq g_j(x_0 + s_k v_k) - g_j(x_0) = s_k \nabla g_j(x_0) \cdot v_k + o(s_k) \quad \forall j \leq \ell' \tag{3.95}$$

Above inequalities are preserved by limit operation, therefore,

$$\nabla f(x_0) \cdot v_k + \frac{o(s_k)}{s_k} \to \nabla f(x_0) \cdot v \leq 0 \tag{3.96}$$
$$\nabla h_i(x_0) \cdot v_k + \frac{o(s_k)}{s_k} \to \nabla h_i(x_0) \cdot v = 0 \tag{3.97}$$
$$\nabla g_j(x_0) \cdot v_k + \frac{o(s_k)}{s_k} \to \nabla g_j(x_0) \cdot v \leq 0 \quad \forall j \leq \ell' \tag{3.98}$$

■

Claim 2: $\nabla g_j(x_0) \cdot v = 0$ for $j = 1, \cdots, \ell''$.

Proof of Claim 2. Suppose not, there exists $j \in \{1, \cdots, \ell''\}$ such that $\nabla g_j(x_0) \cdot v < 0$. Then by (i),

$$0 \geq \nabla f(x_0) \cdot v = -\sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) \cdot v - \sum_{j=1}^{\ell} \mu_j \nabla g_j(x_0) \cdot v \tag{3.99}$$
$$= -\sum_{j=1}^{\ell} \mu_j \nabla g_j(x_0) \cdot v > 0 \tag{3.100}$$

the last inequality is from the fact that $\mu_j \nabla g_j(x_0) \cdot v \leq 0$ for all active constraints and $\mu_j = 0$ for all inactive constraints.

■

(b) and claim 2 suggests $v \in \tilde{\tilde{T}}_{x_0}$.
By the second order Taylor approximation,

$$0 \geq f(x_k) - f(x_0) = s_k \nabla f(x_0) \cdot v_k + \frac{s_k^2}{2} v_k \cdot \nabla^2 f(x_0) \cdot v_k + o(s_k^2) \tag{3.101}$$
$$0 = h_i(x_k) - h_i(x_0) = s_k \nabla h_i(x_0) \cdot v_k + \frac{s_k^2}{2} v_k \cdot \nabla^2 h_i(x_0) \cdot v_k + o(s_k^2) \quad \forall i \tag{3.102}$$
$$0 \geq g(x_k) - g(x_0) = s_k \nabla g_j(x_0) \cdot v_k + \frac{s_k^2}{2} v_k \cdot \nabla^2 g_j(x_0) \cdot v_k + o(s_k^2) \quad \forall j \leq \ell' \tag{3.103}$$

Multiply the second equation by λ_i and third equation by μ_j, and use the fact that $\mu_j = 0$ for every $j > \ell'$. Also, given $\nabla\mathcal{L} = 0$ in (i):

$$0 \geq \frac{s_k^2}{2} v_k \cdot \nabla^2\mathcal{L} \cdot v_k + o(s_k^2) \tag{3.104}$$

Divide by s_k^2 and take the limit $(v_k) \to v$:

$$v \cdot \nabla^2\mathcal{L} \cdot v \leq 0 \tag{3.105}$$

which contradicts the assumption that $\nabla^2\mathcal{L}$ is positive definite in $\tilde{\tilde{T}}_{x_0}$ because we've shown that $v \in \tilde{\tilde{T}}_{x_0}$. ■

4 Iterative Algorithms for Optimization

4.1 Newton's Method

Example 4.1 (Motivation: a second order iterative algorithm)**.** Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ where I is an open interval. Let $x_i \in I$ be a starting point, consider the second order linear approximation of f at x_0:

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \tag{4.1}$$

By construction, the second order Taylor polynomial, $g(x)$, is the best second order approximation to f at x_0 in the following sense:

$$g(x_0) = f(x_0) \tag{4.2}$$
$$g'(x_0) = f'(x_0) \tag{4.3}$$
$$g''(x_0) = f''(x_0) \tag{4.4}$$

The Newton's method aims to solve the critical point of $g(x)$ and define x_1 to be the critical point found:

$$g'(x_1) = f'(x_0) + f''(x_0)(x_1 - x_0) = 0 \tag{4.5}$$
$$\implies x_1 \leftarrow x_0 - \frac{f'(x_0)}{f''(x_0)} \tag{4.6}$$

Algorithm 4.1 (Newton's Method in $\mathbb{R}$)**.** Given initial point $x_0 \in I$, while not terminated:

$$x_{n+1} \leftarrow x_n - \frac{f'(x_n)}{f''(x_n)} \tag{4.7}$$

Theorem 4.1. Let $f \in C^3$ on open interval $I \subseteq \mathbb{R}$. Suppose $x_* \in I$ satisfies $f'(x_*) = 0$ and $f''(x_*) \neq 0$, then the sequence of points (x_n) generated by Newton's method converges to x_* if x_0 *is sufficiently close to* x_*.

Example 4.2. Let $f(x) = x^2$, then $\frac{f'(x)}{f''(x)} = \frac{2x}{2}$. For any starting point x_0, $x_1 = x_0 - \frac{2x_0}{x_0} = 0$. That is, the algorithm converges to the global minimum in one iteration.

Proof. Let $g(x) = f'(x)$ so that $x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}$.
Because $f \in C^3$, then $g \in C^2$.
Note that by $g \in C^2$, $g' = f''$ is bounded away from zero near x_*.
And by continuity again, $g'' = f^{(3)}$ is bounded near the bounded region $V_\varepsilon(x_*)$.

That is, within small region near x_*, $V_\delta(x_*)$, there exists a sufficiently small $\alpha > 0$ such that

$$\begin{cases} |g'(x_1)| > \alpha \ \forall x_1 \in V_\delta(x_*) \\ |g''(x_2)| < \frac{1}{\alpha} \ \forall x_2 \in V_\delta(x_*) \end{cases} \tag{4.8}$$

Further, note that $g(x_*) = f'(x_*) = 0$.
WLOG, let $n \in \mathbb{N}$, suppose $x_n > x_*$:

$$\begin{aligned} x_{n+1} - x_* &= x_n - \frac{g(x_n)}{g'(x_n)} - x_* && (4.9)\\ &= x_n - x_* - \frac{g(x_n) - g(x_*)}{g'(x_n)} && (4.10)\\ &= -\frac{g(x_n) - g(x_*) - g'(x_n)(x_n - x_*)}{g'(x_n)} && (4.11)\\ &= -\frac{1}{2}\frac{g''(\xi)}{g'(x_n)}(x_n - x_*)^2 \quad \text{for some } \xi \in (x_*, x_n) && (4.12) \end{aligned}$$

By taking the absolute values on both sides:

$$\begin{aligned} |x_{n+1} - x_*| &= \frac{1}{2}\frac{|g''(\xi)|}{|g'(x_n)|}|x_n - x_*|^2 && (4.13)\\ &< \frac{1}{2\alpha^2}|x_n - x_*|^2 && (4.14) \end{aligned}$$

Let $\rho := \frac{1}{\alpha^2}|x_0 - x_*|^2$, choose x_0 sufficiently close to x_* such that $\rho < 1$.
Remark: we are showing the iterative algorithm induces a contraction map.
Then,

$$\begin{aligned} |x_1 - x_*| &< \frac{1}{2\alpha^2}|x_0 - x_*|^2 && (4.15)\\ &= \frac{1}{2\alpha^2}|x_0 - x_*|\,|x_0 - x_*| && (4.16)\\ &= \rho\,|x_0 - x_*| && (4.17) \end{aligned}$$

Inductively,

$$\begin{aligned} |x_2 - x_*| &< \frac{1}{2\alpha^2}|x_1 - x_*|^2 && (4.18)\\ &< \frac{1}{2\alpha^2}\rho^2|x_0 - x_*|^2 && (4.19)\\ &= \rho^3|x_0 - x_*| && (4.20)\\ &< \rho^2|x_0 - x_*| && (4.21) \end{aligned}$$

By induction,

$$|x_n - x_*| < \rho^2|x_0 - x_*| \tag{4.22}$$

Therefore, as $n \to \infty$, $(x_n) \to x_*$. ■

Theorem 4.2 (2nd Order MVT)**.**

$$g(x) = g(y) + g'(y)(x-y) + \frac{1}{2}g''(\xi)(x-y)^2 \quad \xi \in (x, y) \tag{4.23}$$

Algorithm 4.2 (Newton's Method in $\mathbb{R}^n$)**.** Let $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ where Ω is open, let initial point $x_0 \in \Omega$. Suppose $\nabla^2 f(x_n)$ is invertible for every generated n, and $\nabla f(x_*) = 0$ so that algorithm stops at minimum. The iterative algorithm is defined as following:

$$x_{n+1} \leftarrow x_n - [\nabla^2 f(x_n)]^{-1} \nabla f(x_n) \tag{4.24}$$

Theorem 4.3 (Generalization)**.** Suppose $x_* \in \Omega$ and $f \in C^3(\Omega, \mathbb{R})$ such that $\nabla f(x_*) = 0$ and $\nabla^2 f(x_*)$ is invertible.TODO: *check this* Then if initial point x_0 is sufficiently closed to x_*, then Newton's method converges to x_*.

Proof. The basic idea is the same as the $\mathbb{R}$ case: prove the iterative algorithm induces a contraction mapping. ∎

Example 4.3 (Newton's Method Fails to Converge)**.** Even if f has an unique global minimum x_*, and x_0 is arbitrarily close to the x_*, Newton's method could fail to converge.
Consider

$$f(x) = \frac{2}{3}|x|^{\frac{3}{2}} \tag{4.25}$$

Note that

$$f(x) = \begin{cases} \frac{2}{3}x^{\frac{3}{2}} & x \geq 0 \\ -\frac{2}{3}x^{\frac{3}{2}} & x < 0 \end{cases} \tag{4.26}$$

$$f'(x) = \begin{cases} x^{\frac{1}{2}} & x \geq 0 \\ -x^{\frac{1}{2}} & x < 0 \end{cases} \tag{4.27}$$

$$f''(x) = \begin{cases} \frac{1}{2}x^{-\frac{1}{2}} & x > 0 \\ -\frac{1}{2}x^{-\frac{1}{2}} & x < 0 \\ \text{DNE} & x = 0 \end{cases} \tag{4.28}$$

Therefore $f \notin C^2$.
Let $\delta > 0$ arbitrarily small, take initial point $x_0 \in V_\delta(0)$. WLOG, $x_0 = \varepsilon \in V_\delta(0)$ with $\varepsilon > 0$. The algorithm will oscillate between $\pm\varepsilon$ and never converge.

Remark 4.1. Newton's method does not necessarily converge to a global minimum, it may converge to local minimum or local maximum or even saddle point.

Example 4.4 (Newton's Method Converges to a Saddle Point)**.** Consider $f(x) = x^3$, $x_{n+1} \to \frac{x_n}{2}$, which converges to 0 (a saddle point).

Example 4.5 (Newton's Method on Quadratic Function)**.** Let Q be a symmetric $n \times n$ invertible matrix. Define quadratic form $f(x) := \frac{1}{2}x^T Q x : \mathbb{R}^n \to \mathbb{R}$. The optimal is $x = 0$.
Let $x_0 \in \mathbb{R}^n$, then $x_1 := x_0 - H_f(x_0)^{-1}\nabla f(x_0) = x_0 - Q^{-1}Qx_0 = 0$. *Therefore, Newton's method converges in one iteration.*

4.2 Steepest/Gradient Descent

Algorithm 4.3 (Steepest Descent)**.** Let $f : \Omega \to \mathbb{R}$ where Ω is an open subset of $\mathbb{R}^n$. Let initial point $x_0 \in \Omega$.
To minimize f on Ω, iteratively update x follows at each step k:

$$x_{k+1} \leftarrow x_k - \alpha_k \nabla f(x_k) \tag{4.29}$$

where $\alpha_k = \operatorname{argmin}_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$.
Remark: There might be multiple minimizing α, in real world implementations, we take the least minimizer found.

Theorem 4.4 (Gradient Descending is Descending)**.** At every step k, if $\nabla f(x_k) = 0$, the algorithm terminates. Otherwise,

$$f(x_{k+1}) < f(x_k) \tag{4.30}$$

Proof. Suppose $\nabla f(x_k) \neq 0$.
Note that for the first minimizing α_k found:

$$\begin{aligned} f(x_{k+1}) &= f(x_k - \alpha_k \nabla f(x_k)) && (4.31)\\ &\leq f(x_k - \alpha \nabla f(x_k)) \quad \forall 0 \leq \alpha \leq \alpha_k && (4.32) \end{aligned}$$

Recall that

$$\left.\frac{d}{ds}\right|_{s=0} f(x_k - s\nabla f(x_k)) = -\nabla f(x_k) \cdot \nabla f(x_k) = -||\nabla f(x_k)||_2^2 < 0 \tag{4.33}$$

Therefore,

$$f(x_{k+1}) \leq f(x_k - \alpha \nabla f(x_k)) < f(x_k) \text{ for small } \alpha \tag{4.34}$$

■

Theorem 4.5 (Gradient Descending Induces Perpendicular Steps)**.** The consecutive steps induced by gradient descending are perpendicular. That is

$$(x_{k+2} - x_{k+1}) \cdot (x_{k+1} - x_k) = 0 \tag{4.35}$$

Proof. Note that

$$\begin{aligned} (x_{k+2} - x_{k+1}) \cdot (x_{k+1} - x_k) &= (-\alpha_{k+1} \nabla f(x_{k+1})) \cdot (-\alpha_k \nabla f(x_k)) && (4.36)\\ &= \alpha_k \alpha_{k+1} \nabla f(x_k) \cdot \nabla f(x_{k+1}) && (4.37) \end{aligned}$$

If $\alpha_k = 0$, done.

If $\alpha_k > 0$,

$$
\begin{aligned}
f(x_{k+1}) &= f(x_k) - \alpha_k \nabla f(x_k) && (4.38)\\
&= \min_{\alpha \geq 0}\{f(x_k - \alpha \nabla f(x_k))\} && (4.39)\\
&= \min_{\alpha > 0}\{f(x_k - \alpha \nabla f(x_k))\} && (4.40)\\
&\implies \left.\frac{\partial}{\partial \alpha}\right|_{\alpha=\alpha_k} f(x_k - \alpha \nabla f(x_k)) = 0 && (4.41)\\
&\implies -\nabla f(x_k - \alpha_k \nabla f(x_k)) \cdot \nabla f(x_k) = 0 && (4.42)\\
&\implies -\nabla f(x_{k+1}) \cdot \nabla f(x_k) = 0 && (4.43)
\end{aligned}
$$

■

Theorem 4.6 (Sufficient Condition for Gradient Descent to Converge)**.** Let $f \in C^1$ on open $\Omega \subseteq \mathbb{R}^n$.
Let $\{x_k\}$ be the sequence generated by gradient descent: $x_{k+1} \leftarrow x_k - \alpha_k \nabla f(x_k)$.
If (x_k) is bounded in Ω, that is, there exists a compact set $K \subseteq \Omega$ such that $(x_k) \subseteq K$,
then every convergent subsequence of (x_k) converges to a critical point $x_* \in \Omega$ of f.

Proof. TODO: *Need to fix this proof.* Let $x_k \in K$ compact.
Then there exists subsequence $x_{k_i} \to x_* \in K$.
Show: $\nabla f(x_*) = 0$.
Note that $f(x_k) \geq f(x_{k+1})$ for every $k \in \mathbb{N}$, therefore $f(x_{k_i}) \searrow f(x_*)$. Therefore, $f(x_k) \searrow f(x_*)$. TODO: *Show this.* Suppose, for contradiction, $\nabla f(x_*) \neq 0$.
By continuity of ∇f, $(\nabla f(x_{k_i})) \to \nabla f(x_*)$.
Let $y_{k_i} := x_{k_i} - \alpha_{k_i} \nabla f(x_{k_i}) = x_{k_{i+1}}$.
Note that y_{k_i} has a convergent subsequence converging to y_*.
WLOG, $(y_{k_i}) \to y_*$.
Observe

$$
\begin{aligned}
\alpha_{k_i} &= \frac{|y_{k_i} - x_{k_i}|}{||\nabla f(x_{k_i})||} && (4.44)\\
\implies \lim_{k_i \to \infty} a_{k_i} &= \frac{|y_* - x_*|}{||\nabla f(x_*)||} =: \alpha_* && (4.45)
\end{aligned}
$$

Put back: $y_* = x_* - \alpha_* \nabla f(x_*)$.
Now $f(y_{k_i}) = f(x_{k_{i+1}}) = \min_{\alpha \geq 0} f(x_{k_i} - \alpha \nabla f(x_{k_i}))$, which implies

$$
\begin{aligned}
& f(y_{k_i}) \leq f(x_{k_i} - \alpha f(x_{k_i}))\ \forall \alpha \geq 0 && (4.46)\\
\forall \alpha \geq 0\ \ & \lim_{i \to \infty} f(y_{k_i}) = f(y_*) \leq \lim_{i \to \infty} f(x_{k_i} - \alpha f(x_{k_i})) = f(x_* - \alpha \nabla f(x_*)) && (4.47)\\
\implies\ & f(y_*) \leq \min_{\alpha \geq 0} f(x_* - \alpha \nabla f(x_*)) < f(x_*) && (4.48)
\end{aligned}
$$

Further note that

$$
f(y_*) = \lim_{i \to \infty} f(y_{k_i}) = \lim_{i \to \infty} f(x_{k_{i+1}}) = f(x_*) \tag{4.49}
$$

Contradiction. ■

4.2.1 Steepest Descent: the Quadratic Case

Example 4.6. Let f follow the general quadratic form

$$f(x) = \frac{1}{2}x^T Qx - b^T x \tag{4.50}$$

with $b, x \in \mathbb{R}^n$ and Q is positive definite.
Let $0 < \lambda = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n = \Lambda$ be eigenvalues of Q.

Proposition 4.1. Gradient descent on strictly convex(concave) quadratic functions is guaranteed to converge to the unique global minimum(maximum).
Essentially, we are going to define an auxiliary function g, preserving the optimizing behaviour in a sense that

$$\operatorname{argmin} g(x) = \operatorname{argmin} f(x) \tag{4.51}$$

and then show the convergence property on $f(x)$ indirectly by showing converging property on $g(x)$.

Lemma 4.1. Recall that positive definite Q implies the existence of unique minimizer x_*. The minimizer satisfies the first order necessary condition.

$$Qx_* - b = 0 \tag{4.52}$$

$$\iff x_* = Q^{-1}b \tag{4.53}$$

Where the second equation came from the invertibility of positive definite matrices. One can rewrite auxiliary function capturing the optimization behaviours f as following

$$g(x) := \frac{1}{2}(x - x_*)^T Q(x - x_*) \tag{4.54}$$

$$= \underbrace{\frac{1}{2}x^T Qx - \overbrace{x^T Qx_*}^{x^T b}}_{f(x)} + \frac{1}{2}x_*^T Qx_* \tag{4.55}$$

Because Q is positive definite:

$$g(x) \geq 0 \ \forall x \in \mathbb{R}^n \tag{4.56}$$

$$g(x) = 0 \iff x = x_* \tag{4.57}$$

$$\nabla f(x) = \nabla g(x) = Qx - b \tag{4.58}$$

Where the last equation came from the fact that $f(x)$ and $g(x)$ differ by a constant.
Therefore, using the method of steepest descent,

$$x_{k+1} = x_k - \alpha_k \nabla g(x_k) \tag{4.59}$$

where $\alpha_k \in \operatorname{argmin}_{\alpha \geq 0} f(x_k - \alpha \nabla g(x_k))$.

The necessary condition for minimizations suggests α_k must satisfy

$$
\begin{aligned}
0 = \left.\frac{d}{d\alpha}\right|_{\alpha=\alpha_k} f(x_k - \alpha\nabla g(x_k)) &= \nabla f(x_k - \alpha_k \nabla g(x_k)) \cdot (-\nabla g(x_k)) && (4.60)\\
&= -[Q(x_k - \alpha_k \nabla g(x_k)) - b] \cdot \nabla g(x_k) && (4.61)\\
&= -[Qx_k - \alpha_k Q\nabla g(x_k) - b] \cdot \nabla g(x_k) && (4.62)\\
&= -[\nabla g(x_k) - \alpha_k Q\nabla g(x_k)] \cdot \nabla g(x_k) && (4.63)\\
&= -||\nabla g(x_k)||_2^2 + \alpha_k \nabla g(x_k)^T Q \nabla g(x_k) && (4.64)\\
\implies \alpha_k &= \frac{||\nabla g(x_k)||_2^2}{\nabla g(x_k)^T Q \nabla g(x_k)} && (4.65)
\end{aligned}
$$

Assumption 4.1. TODO: *Need to check if this assumption is required.*
Assume Q is symmetric.

Lemma 4.2. The iterative updating from gradient descent on $g(x)$ is a contraction mapping. That is,

$$
g(x_{k+1}) = \underbrace{\left(1 - \frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-1} \nabla g(x_k)]}\right)}_{\in[-1,1]} g(x_k) \tag{4.66}
$$

Proof.

$$
\begin{aligned}
g(x_{k+1}) &\equiv g(x_k - \alpha_k \nabla g(x_k)) && (4.67)\\
&\equiv \frac{1}{2}[x_k - \alpha_k \nabla g(x_k) - x_*]^T Q[x_k - \alpha_k \nabla g(x_k) - x_*] && (4.68)\\
&= \frac{1}{2}[x_k - x_* - \alpha_k \nabla g(x_k)]^T Q[x_k - x_* - \alpha_k \nabla g(x_k)] && (4.69)\\
&= \underbrace{\frac{1}{2}(x_k - x_*)^T Q(x_k - x_*)}_{g(x_k)} - \alpha_k \nabla g(x_k)^T Q(x_k - x_*) + \frac{1}{2} a_k^2 \nabla g(x_k)^T Q \nabla g(x_k) && (4.70)
\end{aligned}
$$

$$
\implies g(x_k) - g(x_{k+1}) = -\frac{1}{2} a_k^2 \nabla g(x_k)^T Q \nabla g(x_k) + \alpha_k \nabla g(x_k)^T Q \underbrace{(x_k - x_*)}_{=:y_k} \tag{4.71}
$$

$$
\begin{aligned}
\implies \frac{g(x_k) - g(x_{k+1})}{g(x_k)} &= \frac{-\frac{1}{2} a_k^2 \nabla g(x_k)^T Q \nabla g(x_k) + \alpha_k \nabla g(x_k)^T Q y_k}{\frac{1}{2} y_k^T Q y_k} && (4.72)\\
&= \frac{2\alpha_k \nabla g(x_k)^T Q y_k - a_k^2 \nabla g(x_k)^T Q \nabla g(x_k)}{y_k^T Q y_k} && (4.73)
\end{aligned}
$$

Note that the first order condition implies $Qx_* = b$.
Therefore, $\nabla g(x_k) = Qx_k - b = Qx_k - Qx_* = Qy_k$, which implies $y_k = Q^{-1}\nabla g(x_k)$.

$$
\begin{aligned}
\frac{2\alpha_k \nabla g(x_k)^T Q y_k - a_k^2 \nabla g(x_k)^T Q \nabla g(x_k)}{y_k^T Q y_k} &= \frac{2\alpha_k \nabla g(x_k)^T Q Q^{-1} \nabla g(x_k) - a_k^2 \nabla g(x_k)^T Q \nabla g(x_k)}{\nabla g(x_k)^T Q^{-T} Q Q^{-1} \nabla g(x_k)} && (4.74)\\
&= \frac{2\alpha_k ||\nabla g(x_k)||_2^2 - a_k^2 \nabla g(x_k)^T Q \nabla g(x_k)}{\nabla g(x_k)^T Q^{-T} \nabla g(x_k)} && (4.75)
\end{aligned}
$$

Plug in the α_k computed before:

$$\dots = \frac{2\frac{||\nabla g(x_k)||_2^2}{\nabla g(x_k)^T Q \nabla g(x_k)}||\nabla g(x_k)||_2^2 - \frac{||\nabla g(x_k)||_2^4}{(\nabla g(x_k)^T Q \nabla g(x_k))^2}\nabla g(x_k)^T Q \nabla g(x_k)}{\nabla g(x_k)^T Q^{-T} \nabla g(x_k)} \tag{4.76}$$

$$= \frac{2\frac{||\nabla g(x_k)||_2^4}{\nabla g(x_k)^T Q \nabla g(x_k)} - \frac{||\nabla g(x_k)||_2^4}{\nabla g(x_k)^T Q \nabla g(x_k)}}{\nabla g(x_k)^T Q^{-T} \nabla g(x_k)} \tag{4.77}$$

$$= \frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-T} \nabla g(x_k)]} \tag{4.78}$$

$$= \frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-1} \nabla g(x_k)]} \quad \because Q \in \mathbb{S}_n \tag{4.79}$$

$$\implies g(x_k) - g(x_{k+1}) = \left\{\frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-1} \nabla g(x_k)]}\right\} g(x_k) \tag{4.80}$$

$$\implies g(x_{k+1}) = \left\{1 - \left[\frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-1} \nabla g(x_k)]}\right]\right\} g(x_k) \tag{4.81}$$

Lemma 4.3 (Kantorovich Inequality)**.** Let Q be a $n \times n$ positive definite symmetric matrix with eigenvalues $0 < \lambda = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n = \Lambda$. Then, for any $v \in \mathbb{R}^n$:

$$\frac{||v||_2^4}{(v^T Q v)(v^T Q^{-1} v)} \geq \frac{4\lambda\Lambda}{(\lambda + \Lambda)^2} \tag{4.82}$$

Proof of Kantorovich Inequality. TODO: *Prove this lemma* ∎

Therefore,

$$g(x_{k+1}) = \left\{1 - \left[\frac{||\nabla g(x_k)||_2^4}{[\nabla g(x_k)^T Q \nabla g(x_k)][\nabla g(x_k)^T Q^{-1} \nabla g(x_k)]}\right]\right\} g(x_k) \tag{4.83}$$

$$\leq \left\{1 - \frac{4\lambda\Lambda}{(\lambda + \Lambda)^2}\right\} g(x_k) \tag{4.84}$$

$$= \frac{(\lambda - \Lambda)^2}{(\lambda + \Lambda)^2} g(x_k) \tag{4.85}$$

∎

Theorem 4.7. For any initial point $x_0 \in \mathbb{R}^n$, gradient descent converges to the unique minimum point x_* of the quadratic $f(x) = x^T Q x - b^T x$.

Proof. Define $q(x) := \frac{1}{2}(x - x_*)^T Q(x)(x - x_*)$.
Note that $q(x)$ and $f(x)$ differ by a constant, therefore $\operatorname{argmin} q(x) = \operatorname{argmin} f(x)$.
Moreover, we've shown:

$$q(x_{k+1}) \leq \underbrace{\left(\frac{\Lambda - \lambda}{\Lambda + \lambda}\right)^2}_{\in[0,1)} q(x_k) \tag{4.86}$$

It is easy to notice that

$$q(x_k) \leq r q(x_{k-1}) \tag{4.87}$$

$$\implies q(x_k) \leq r^k q(x_0) \tag{4.88}$$

$$\implies q(x_k) \in \{x \in \mathbb{R}^k : q(x) \leq r^k q(x_0)\} =: \mathcal{L}_k \tag{4.89}$$

Note that the sub-level set $\mathcal{L}_k$ is strictly decreasing (i.e. $\mathcal{L}_{k+1} \subsetneq \mathcal{L}_k$).
Further, note that x_* is the only point satisfying the inequality at the limit:

$$q(x_*) = 0 = \lim_{k\to\infty} q(x_0) \tag{4.90}$$

Therefore, $\lim_{k\to\infty} \mathcal{L} = \{0\}$, and $(x_k) \to x_*$. ■

Remark 4.2. Note that

$$r = \left(\frac{\frac{\Lambda}{\lambda} - 1}{\frac{\Lambda}{\lambda} + 1}\right)^2 \in [0, 1) \tag{4.91}$$

$$= \left(\frac{C - 1}{C + 1}\right)^2 \tag{4.92}$$

where $C = \frac{\Lambda}{\lambda}$ is the **condition number** of Q.
Clearly, when $\lambda = \Lambda$, $r = 0$ and gradient descent converges to the unique global minimum after only one epoch.
While $C \gg 1$, $r \approx 1$ and the worst case of gradient descent converges slowly.