ISyE 6663 (Winter 2017)

Nonlinear Optimization

Prof. R. D. C. Monteiro
Georgia Institute of Technology

LaTeXer: W. Kong
http://wwkong.github.io
Last Revision: May 3, 2017

Table of Contents

These notes are currently a work in progress, and as such may be incomplete or contain errors.

Acknowledgments:

Special thanks to *Michael Baker* and his LaTeX formatted notes. They were the inspiration for the structure of these notes.

Abstract

The purpose of these notes is to provide the reader with a secondary reference to the material covered in ISyE 6663.

1 Review of Concepts

1.1 Unconstrained Optimization

Definition 1.1. For a set $S \subseteq \mathbb{R}^n$ and $f : S \mapsto \mathbb{R}$, an optimization problem can be formulated as

$$\begin{aligned} \min(\max)\ & f(x) \\ \text{s.t. } & x \in S \end{aligned}$$

which we will call the standard minimization problem.

Example 1.1. Here are some basic examples:

(1) $S = \{x \in \mathbb{R}^n : Ax \le b\}$

(2) $S = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \le 0, x \in X\}$ where $h : \mathbb{R}^n \mapsto \mathbb{R}^m, g : \mathbb{R}^n \mapsto \mathbb{R}^r$, and $X \subseteq \mathbb{R}^n$ is simple

Remark 1.1. If $S = \mathbb{R}^n$ then the problem is unconstrained, otherwise if $S \neq \mathbb{R}^n$ then it is constrained.

Example 1.2. (Least Squares) For error $e_i = y_i - \hat{f}(t_i)$ and trial function $\hat{f}(t) = x_1 + x_2 \exp(-x_3 t)$, a constrained optimization problem is

$$\begin{aligned} \min\ & \sum_{i=1}^m e_i^2 = \sum_{i=1}^m (y_i - x_1 - x_2 e^{-x_3 t_i}) \\ \text{s.t. } & x_3 \ge 0 \\ & x_1 + x_2 = 1 \end{aligned}$$

Definition 1.2. $x^* \in S$ is a [strict] **global minimum** (optimal solution) of the standard minimization problem if $f(x)\ [>]\ \ge f(x^*)\ [x \neq x^*]$ for all $x \in S$. Similar definitions follow for maximization problems.

Notation. We will denote:

$$\begin{aligned} B(x^*; \varepsilon) &= \{x \in \mathbb{R}^n : \|x - x^*\| < \varepsilon\} \\ \bar{B}(x^*; \varepsilon) &= \{x \in \mathbb{R}^n : \|x - x^*\| \le \varepsilon\} \end{aligned}$$

Definition 1.3. $x^* \in S$ is a [strict] **local minimum** of the standard minimization problem $\exists \varepsilon > 0$ such that $f(x)\ [>]\ \ge f(x^*)$ $[x \neq x^*]$ for all $x \in S \cap \bar{B}(x^*, \varepsilon)$.

Definition 1.4. S is compact iff S is closed and bounded.

Theorem 1.1. *(Weierstrass) If S is compact and f is continuous on S, then the standard minimization problem has a global minimum.*

Corollary 1.1. *If S is closed and f is continuous on S and $\lim_{\|x\| \to \infty, x \in S} f(x) = \infty$ then the standard minimization problem has a global minimum. The condition $\lim_{\|x\| \to \infty, x \in S} f(x) = -\infty$ is instead required for maximization problems.*

Note that:

$$\begin{aligned} \lim_{\|x\| \to \infty, x \in S} f(x) = \infty &\iff (\forall M \ge 0, \exists r \ge 0 \text{ s.t. } \|x\| > r, x \in S \implies f(x) > M) \\ &\iff (\forall M \ge 0, \exists r \ge 0 \text{ s.t. } f(x) \le M \implies x \in S, \|x\| \le r) \\ &\iff \{x \in S : f(x) \le M\} \subseteq \bar{B}(0, r) \\ &\iff \forall M \ge 0, \{x \in S : f(x) \le M\} \text{ is bounded.} \end{aligned}$$

Proof. (Sketch) Pick $x_0 \in S$ such that $M = f(x_0)$ and remark that $\{x \in S : f(x) \le f(x_0)\}$ is compact. The rest follows from Weierstrass. □

Definition 1.5. Given $S = \mathbb{R}^n, f : \mathbb{R}^n \mapsto \mathbb{R}, \bar{x} \in \mathbb{R}^n$, the **gradient** of f at $\bar{x}$ is

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), ..., \frac{\partial f}{\partial x_n}(\bar{x}) \right)^T \in \mathbb{R}^n$$

Remark 1.2. (Interpretations)

(1) In the set $\{x \in \mathbb{R}^n : f(x) = f(\bar{x})\}$ the gradient lies perpendicular to this set and points in the direction of steepest ascent.

(2) The **graph** of the function f is $\{(x, f(x)) \in \mathbb{R}^{n+1} : x \in \mathbb{R}^n\}$ and the gradient defines a linear approximation at $\bar{x}$ given by $t = f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle$. In particular,

$$0 = \begin{pmatrix} -\nabla f(\bar{x}) \\ 1 \end{pmatrix}^T \begin{pmatrix} x - \bar{x} \\ t - f(\bar{x}) \end{pmatrix}$$

Proposition 1.1. *x^* is a local minimum of the standard optimization problem and f is differentiable at x^* $\implies \nabla f(x^*) = 0$.*

Proof. Let $d \in \mathbb{R}^n$ be given. For every $t > 0$ sufficiently small, $0 \leq \frac{f(x^*+td)-f(x)}{t}$ as $t \to 0^+$ we get $\langle \nabla f(x^*), d \rangle = \nabla f(x^*)^T d \geq 0$ for any $d \in \mathbb{R}^n$. This is only the case for when $\nabla f(x^*) = 0$ as the case for $d = -\nabla f(x^*) \implies -\|\nabla f(x^*)\|^2 \geq 0$. □

Definition 1.6. $H \in \mathbb{R}^{n \times n}$ is **positive semi-definite** if $x^T H x \geq 0$ for all $x \in \mathbb{R}^n$ (Notation $H \succeq 0$). It is **positive definite** if $x^T H x > 0$ for all $x \in \mathbb{R}^n, x \neq 0$.

Fact 1.1. *If f is twice continuously differentiable at x, then*

$$\nabla^2 f(x) = f''(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right]_{ij}$$

is symmetric.

Proposition 1.2. *x^* is a local minimum of the standard optimization problem and f is twice continuously differentiable at x^* (or $f \in \mathcal{C}^2(\mathbb{R})$) $\implies \nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \geq 0$.*

Proof. Note that

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x) h + r(h)\|h\|^2$$
$$\lim_{\|h\| \to 0} r(h) = 0$$

or equivalently

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x + th) h$$

for some $t \in (0,1)$. The case for $\nabla f(x^*) = 0$ has already been shown so let $H = \nabla^2 f(x^*)$ and $d \in \mathbb{R}^n$. We want to show that $d^T H d \geq 0$. We have for $t > 0$ sufficiently small,

$$0 \leq f(x^* + td) - f(x^*) = \underbrace{t \nabla f(x)^T d}_{=0} + \frac{1}{2} t^2 d^T H d + t^2 r(td) \|d\|^2$$

from the first expansion. Dividing by t^2 gives us

$$0 \leq \frac{1}{2} d^T H d + r(td) \|d\|^2.$$

Taking $t \to 0$ yields $0 \leq d^T H d$. □

Example 1.3. The converse is generally not true. Consider the case $f(x) = x^3$ which satisfies the first and second order conditions at $x = 0$ but does not have a local minimum at that point.

Theorem 1.2. *Assume that $f \in \mathcal{C}^2$ and $x^* \in \mathbb{R}^n$ is such that $\nabla f(x^*) = 0, \nabla^2 f(x^*) > 0$. Then x^* is a strict local minimizer of the standard minimization problem.*

Proof. Let $H = \nabla^2 f(x^*)$. By Weierstrass Theorem, choose $\alpha > 0$ such that $u^T H u \geq \alpha$ for all $u \in \mathbb{R}^n$ such that $\|u\| \leq 1$. We have

$$f(x^* + h) - f(x^*) = \frac{1}{2} h^T H h + r(h) \|h\|^2$$
$$\lim_{\|h\| \to 0} r(h) = 0$$

which implies that $\exists\delta > 0$ such that $\|h\| \le \delta \implies |r(h)| \le \frac{\alpha}{4}$ and hence, if $\|h\| \le \delta$, we have

$$\begin{aligned} f(x^* + h) - f(x^*) &= \|h\|^2 \left[\frac{1}{2}\left(\frac{h^T}{\|h\|}\right) H \left(\frac{h}{\|h\|}\right) + r(h)\right] \\ &\ge \|h\|^2 \left[\frac{\alpha}{2} - \frac{\alpha}{4}\right] = \frac{1}{4}\alpha\|h\|^2 \end{aligned}$$

Hence, if $0 < \|h\| \le \delta$ then $f(x^* + h) - f(x^*) > 0$. So, x^* is a local minimum of the standard minimization problem. □

Example 1.4. The above condition is not necessary and the converse is not true. Consider the function $f(x) = x^4$ at $x^* = 0$.

1.2 Convexity

Definition 1.7. $C \subseteq \mathbb{R}^n$ is a **convex set** if $(x, y) := \{tx + (1-t)y : t \in (0,1)\} \subseteq C$ for all $x, y \in C$. Here are some properties:

1) If $\{C_i\}_{i\in I}$ is a collection of convex sets in $\mathbb{R}^n$ then $\bigcap_{i\in I} C_i$ is convex.

2) If $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is affine, $C \subseteq \mathbb{R}^n$, and $D \subseteq \mathbb{R}^m$ then $T(C), T^{-1}(D)$ are convex.

3) $C_i \subseteq \mathbb{R}^{n_i}$ convex for $i = 1, 2, ..., r$ implies that $C_1 \times ... \times C_r$ is convex

4) $C_i \subseteq \mathbb{R}^n$ convex for $i = 1, 2, ..., r$ implies that $C_1 + ... + C_r$ (Minkowski sum) is convex

5) $C \subseteq \mathbb{R}^n$ is convex, $\alpha \in \mathbb{R}$ implies that αC is convex

6) C convex implies that cl(C) and int(C) are convex

Example 1.5. Here are some examples:

1) Hyperplane: $0 \ne u \in \mathbb{R}^n, \beta \in \mathbb{R}$ define $H = H(u, \beta) := \{x \in \mathbb{R}^n : u^T x = \beta\}$

2) Half-spaces: $H^+ = \{x \in \mathbb{R}^n : u^T x \ge \beta\}, H^- = \{x \in \mathbb{R}^n : u^T x \le \beta\}$

3) Polyhedra: $\bigcap_i H_i^-$

Proposition 1.3. *If C is convex then $\sum_{i=1}^n \alpha_i x^i \in C$ for $x^i \in C, \alpha_i \ge 0, i = 1, 2, ..., n$, and $\sum_{i=1}^n \alpha_i = 1$.*

Proof. (Can be done by induction, using convexity) □

Definition 1.8. Let $C \subseteq \mathbb{R}^n$ be a convex set and $f(x)$ be a unction defined on C. A function f is **convex** on C if for all $x, y \in C, t \in (0,1)$ we have

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y)$$

The function f is **strictly convex** if the above inequality if the above holds strictly whenever $x \ne y$.

Definition 1.9. f is **β-strongly convex** ($\beta > 0$) on C if

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) - \frac{\beta}{2}t(1-t)\|x - y\|^2$$

for all $x, y \in C$ and $t \in (0,1)$.

Proposition 1.4. *f is β-strongly convex iff $f - \frac{\beta}{2}\|\cdot\|^2$ is strongly convex.*

Proof. (Left as an exercise) □

Proposition 1.5. *If $f$is convex on C then for every $\alpha \in \mathbb{R}$, the sets*

$$\{x \in C : f(x) < \alpha\}$$
$$\{x \in C : f(x) \le \alpha\}$$

are convex.

Proposition 1.6. *The following are equivalent*

(1) f *is convex on* C

(2) $\{x, t \in C \times \mathbb{R} : f(x) \leq t\}$ *is convex*

(2) $\{x, t \in C \times \mathbb{R} : f(x) < t\}$ *is convex*

Proof. Left as an exercise to the reader. □

Proposition 1.7. *(Jensen's inequality) Assume* f *is convex on* C*. If* $x^1, ..., x^r \in C$ *with* $\sum_{i=1}^r \alpha_i = 1, \alpha_1, ..., \alpha_r \geq 0$ *then* $f\left(\sum_{i=1}^r \alpha_i x^i\right) \leq \sum_{i=1}^r \alpha_i f(x^i)$.

Proof. Since f is convex, the set

$$U = \{(x, t) \in C \times \mathbb{R} : f(x) \leq t\}$$

is convex. Clearly, $(x^i, f(x^i))^T \in U$ for all $i = 1, ..., r$. So $\sum_i \alpha_i (x^i, f(x^i))^T = (\sum_i \alpha_i x^i, \sum_i \alpha_i f(x^i))^T \in U$ and hence $f\left(\sum_{i=1}^r \alpha_i x^i\right) \leq \sum_{i=1}^r \alpha_i f(x^i)$ from the definition of U. □

Notation 1. For $\Omega \subseteq \mathbb{R}^n$ we write $f \in \mathcal{C}^1(\Omega)$ if f is continuously differentiable of every $x \in \Omega$.

Proposition 1.8. *For* $\Omega \subseteq \mathbb{R}^n$ *convex and* $f \in \mathcal{C}^1(\Omega)$ *the following are equivalent:*

(a) f *is (strictly) convex on* Ω

(b) $f(y)(>) \geq f(x) + \nabla f(x)^T (y - x), \forall x, y \in \Omega$ *(*$x \neq y$*)*

(c) $[\nabla f(y) - \nabla f(x)]^T (y - x)(>) \geq 0, \forall x, y \in \Omega$ *(*$x \neq y$*)*

Proof. $[(a) \implies (b)]$ Let $x, y \in \Omega$ be given. For all $t \in (0, 1)$, we have

$$\begin{aligned}
& f(ty + (1-t)x) \leq t f(y) + (1-t) f(x) \\
\implies & f(x + t(y - x)) \leq f(x) + t[f(y) - f(x)] \\
\implies & \frac{f(x + t(y-x)) - f(x)}{t} \leq f(y) - f(x) \\
\overset{t \to 0}{\implies} & \nabla f(x)^T (y - x) \leq f(y) - f(x)
\end{aligned}$$

□

$[(b) \implies (a)]$ Let $x, y \in \Omega$ be given and $z_t = ty + (1-t)x \in \Omega$. Then by (b),

$$\begin{cases} f(y) \geq f(z_t) + \nabla f(z_t)^T (y - z_t) & (1) \\ f(x) \geq f(z_t) + \nabla f(z_t)^T (x - z_t) & (2) \end{cases}$$

and $(t)(1) + (1-t)(2)$ yields

$$\begin{aligned}
t f(y) + (1-t) f(x) &\geq f(z_t) + \nabla f(z_t)^T (ty + (1-t)x - z_t) \\
&= f(z_t) = f(ty + (1-t)x)
\end{aligned}$$

$[(b) \implies (c)]$ Just add the two inequalities:

$$\begin{cases} f(y) \geq f(x) + \nabla f(z_t)^T (y - x) \\ f(x) \geq f(y) + \nabla f(z_t)^T (x - y) \end{cases}$$

$[(c) \implies (b)]$ For some $t \in (0, 1)$,

$$f(y) - f(x) = \nabla f(z_t)^T (y - x)$$

where $z_t = x + t(y - x)$. Since $z_t - x = t(y - x)$ we have

$$\begin{aligned} & f(y) - f(x) - \nabla f(x)^T (y - x) \\ = & \left[\nabla f(z_t) - \nabla f(x)\right]^T (y - x) \\ = & \frac{1}{t} \left[\nabla f(z_t) - \nabla f(x)\right]^T (z_t - x) \geq 0 \end{aligned}$$

Proposition 1.9. *f is β-strongly convex ($\beta > 0$) iff $f - \frac{\beta\|\cdot\|^2}{2}$ is convex.*

Proposition 1.10. *For $\Omega \subseteq \mathbb{R}^n$ convex, $f \in \mathcal{C}^1(\Omega)$ and $\beta \in \mathbb{R}$, the following are equivalent:*

(a) $f - \frac{\beta\|\cdot\|^2}{2}$ is convex

(b) $\forall x, y \in \Omega$, $f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\beta}{2}\|y - x\|^2$

(c)$\forall x, y \in \Omega$, $[\nabla f(y) - \nabla f(x)]^T (y - x) \geq \beta\|y - x\|^2$

Proof. Define $\tilde{f} := f - \frac{\beta\|\cdot\|^2}{2}$ and remark that $\tilde{f}$ is convex and from a previous result,

$$\begin{aligned} & \nabla \tilde{f}(x) = \nabla f(x) - \beta x \\ \iff & \tilde{f}(y) \geq \tilde{f}(x) + \nabla \tilde{f}(x)^T (y - x), \forall x, y \in \Omega \qquad (1) \\ \iff & [\nabla \tilde{f}(y) - \nabla \tilde{f}(x)]^T (y - x) \geq 0, \forall x, y \in \Omega \qquad (2) \end{aligned}$$

and (1) is equivalent to (b) and (2) is equivalent to (c). □

Proposition 1.11. *For $\Omega \subseteq \mathbb{R}^n$ convex, $f \in \mathcal{C}^1(\Omega)$ and $M \in \mathbb{R}$, the following are equivalent:*

(a) $\frac{M}{2}\|\cdot\| - f$ is convex

(b) $\forall x, y \in \Omega$, $f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{M}{2}\|y - x\|^2$

(c)$\forall x, y \in \Omega$, $[\nabla f(y) - \nabla f(x)]^T (y - x) \leq M\|y - x\|^2$

Proof. Apply the previous proposition with $\beta = -M, f = -f$. □

Proposition 1.12. *Assume $\Omega \subseteq \mathbb{R}^n$ is convex, $f \in \mathcal{C}^1(\Omega)$ is convex on Ω. The the following are equivalent for $\bar{x} \in \mathbb{R}^n$:*

(a) $\bar{x}$ is a global minimum of f on Ω

(b) $\bar{x}$ is a local minimum of f on Ω

(c) $\nabla f(\bar{x})^T (x - \bar{x}) \geq 0$, $\forall x \in \Omega$ or $f'(\bar{x}; x - \bar{x}) \geq 0$ where $f'(\bar{x}; x - \bar{x}) = \nabla f(\bar{x})^T (x - \bar{x})$

Proof. $[(a) \implies (b)]$ Obvious.

$[(b) \implies (c)]$ Since $\bar{x}$ is a local minimum, $f(\bar{x} + t(x - \bar{x})) - f(\bar{x}) \geq 0$ for $t > 0$ sufficiently small. If we divide by t and take $t \to 0$ then $\nabla f(\bar{x})^T (\bar{x} - x)$.

$[(c) \implies (a)]$ Let $x \in \Omega$ be given. By (c), $\nabla f(\bar{x})^T (x - \bar{x}) \geq 0$. By the convexity of f, we have

$$\begin{aligned} & f(x) \geq f(\bar{x}) + \nabla f(x)^T (x - \bar{x}) \\ \implies & f(x) \geq f(\bar{x}) \end{aligned}$$

and so $\bar{x}$ is a global minimum. □

Remark 1.3. If $\bar{x} \in \text{int}(\Omega)$ then $(c) \iff \nabla f(\bar{x}) = 0$.

Proof. $(\implies)$ Assume $\nabla f(\bar{x}) \neq 0$. We know $\exists \epsilon > 0$ such that $\bar{B}(\bar{x}; \epsilon) \subseteq \Omega$ and

$$\nabla f(\bar{x})^T (x - \bar{x}) \geq 0, \forall x \in \Omega, \forall x \in \bar{B}(\bar{x}; \epsilon)$$

from (c). Now $x := x - \epsilon \frac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|} \in \bar{B}(\bar{x}; \epsilon)$ and substituting this into the above equation yields $0 \leq -\epsilon\|\nabla f(\bar{x})\| < 0$ leading to a contradiction. □

Proposition 1.13. *If $\Omega \subseteq \mathbb{R}^n$ is convex, $f \in \mathcal{C}^1(\Omega)$ is strictly convex on Ω then f has at most one global minimum.*

Proof. Assume $\bar{x} \in \Omega$ is a global minimum of $\min\{f(x) : x \in \Omega\}$. Let $x \neq \bar{x}, x \in \Omega$ be given. We have

$$f(x) > f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$$

and $\nabla f(\bar{x})^T (x - \bar{x}) \geq 0$ from a previous result. So $f(x) > f(\bar{x})$ and thus $\bar{x}$ is the only global minimum. □

Proposition 1.14. *If Ω is convex, $f \in \mathcal{C}^1(\Omega)$, $\nabla f(\cdot)$ is $L-$Lipschitz continuous on Ω (i.e. $\|\nabla f(y) - \nabla f(x)\| \leq L\|x - y\|$ for all $x, y \in \Omega$), then*

$$-\frac{L}{2}\|x - y\|^2 \leq f(y) - [f(x) + \nabla f(x)^T (y - x)] \leq \frac{L}{2}\|x - y\|^2,$$
$$-L\|x - y\|^2 \leq [\nabla f(y) - \nabla f(x)]^T (y - x) \leq L\|x - y\|^2.$$

The second set of inequalities is proven by Cauchy-Schwarz.

Proposition 1.15. *If $\Omega \subseteq \mathbb{R}^n$ is closed and convex, and $f \in \mathcal{C}^1(\Omega)$ is $\beta-$strongly convex. Then,*

$$f_* = \inf_x \{f(x) : x \in \Omega\}$$

has a unique optimal solution x^* and

$$f(x) \geq f_* + \frac{\beta}{2}\|x - x^*\|^2, \forall x \in \Omega$$

Proof. Take $x_0 \in \Omega$. Since f is $\beta-$strongly convex, we have

$$f(x) \geq f(x_0) + \nabla f(x)^T (x - x_0) + \frac{\beta}{2}\|x - x_0\|^2$$

for all $x \in \Omega$. Hence, as $\|x\| \to \infty, x \in \Omega$, we will have $f(x) \to \infty$. Thus, $\inf\{f(x) : x \in \Omega\}$ has a unique optimal solution x^*. Hence, $\nabla f(x^*)^T (x - x^*) \geq 0$ for all $x \in \Omega$ and

$$f(x) \geq f_* + \frac{\beta}{2}\|x - x^*\|^2, \forall x \in \Omega.$$

□

1.3 Projection onto Convex Sets

Definition 1.10. For $\Omega \subseteq \mathbb{R}^n$ closed and convex, $x \in \mathbb{R}^n$, we define

$$\Pi_\Omega(x) = \underset{y}{\operatorname{argmin}}\{\|y - x\| : y \in \Omega\} = \underset{y}{\operatorname{argmin}}\left\{\frac{1}{2}\|y - x\|^2 : y \in \Omega\right\}$$

as the **projection** of x onto Ω. The latter definition is useful because the $\frac{1}{2}\|\cdot\|$ function is strongly convex.

Corollary 1.2. *Using the previous definition and $\langle x, y\rangle \equiv x^T y$,*

(1) Π_Ω is well-defined

(2) $x^* = \Pi_\Omega(x) \iff \langle y - x^*, x - x^*\rangle \leq 0, \forall y \in \Omega$

(3) $\langle x_1 - x_2, \Pi_\Omega(x_1) - \Pi_\Omega(x_2)\rangle \geq \|\Pi_\Omega(x_1) - \Pi_\Omega(x_2)\|^2$ and hence $\|x_1 - x_2\| \geq \|\Pi_\Omega(x_1) - \Pi_\Omega(x_2)\|, \forall x_1, x_2 \in \Omega$. That is, Π_Ω is non-expansive.

Proof. (1) is obvious. For (2), let $f(y) = \frac{1}{2}\|y - x\|^2$. Then,

$$\begin{aligned}
&x^* = \Pi_\Omega(x) \\
\iff &x^* \in \underset{y}{\operatorname{argmin}}\{f(y) : y \in \Omega\} \\
\iff &\nabla f(x^*)^T(y - x^*) \geq 0, \forall y \in \Omega \\
\iff &(x^* - y)^T(y - x^*) \geq 0, \forall y \in \Omega.
\end{aligned}$$

For (3), define $x_i^* = \Pi_\Omega(x_i), i = 1, 2$. We have

$$\begin{aligned}
(x_1 - x_1^*)^T(x_2 - x_2^*) \leq 0 \\
(x_2 - x_2^*)^T(x_1 - x_1^*) \leq 0
\end{aligned}$$

and adding the two above inequalities yields

$$\begin{aligned}
&[(x_1 - x_2) - (x_1^* - x_2^*)]^T(x_2^* - x_1^*) \leq 0 \\
\implies &\|x_1^* - x_2^*\|^2 \leq (x_2 - x_1)^T(x_2^* - x_1^*) \leq \|x_2 - x_1\|\|x_2^* - x_1^*\|.
\end{aligned}$$

□

Remark 1.4. If Ω is closed convex, $\bar{x} \in \Omega$, and we define the **normal cone** of $\bar{x}$ as

$$N_\Omega(\bar{x}) = \{n \in \mathbb{R}^n : n^T(y - \bar{x}) \leq 0, y \in \Omega\},$$

then the second condition of the previous propositions says $0 \in x^* + N_\Omega(x^*) - x$.

Remark 1.5. If f is convex [I'm assuming we need this], then the problem $\min_y \{f(y) : y \in \Omega\}$ is equivalent to $0 \in \nabla f(x^*) + N_\Omega(x^*)$. This follows from the fact that the optimality condition for the problem is

$$\nabla f(x^*)^T(y - x^*) \geq 0, \forall y \in \Omega \iff -\nabla f(x^*) \in N_\Omega(x^*).$$

Proposition 1.16. *Assume $\Omega \subseteq \mathbb{R}^n$ convex and $f \in \mathcal{C}^1(\Omega)$. Then,*

(a) $\nabla^2 f(x) \geq 0, \forall x \in \Omega \implies f$ is convex on Ω.

(b) f is convex on Ω and int $\Omega \neq \emptyset \implies \nabla^2 f(x) \geq 0, \forall x \in \Omega$.

(c) $\nabla^2 f(x) > 0, \forall x \in \Omega \implies f$ is strictly convex on Ω.

Proof. (a) Let $x, y \in \Omega$. We will show $f(y) \geq f(x) + \nabla f(x)^T(y - x)$. We have

$$f(y) = f(x) + \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T\nabla^2 f(\xi)(y - x)$$

for some $\xi = x + t(x - y)$ and $t \in (0, 1)$. Clearly $\xi \in \Omega$ and hence $\nabla^2 f(\xi) \geq 0$. So $d^T\nabla^2 f(\xi)d \geq 0, \forall d \in \mathbb{R}^n$ and the result follows.

(b) By contradiction, assume $\exists x \in \Omega$ such that $\nabla^2 f(x) \ngeq 0$. Without loss of generality, we may assume that $x \in \text{int}\,\Omega$ from the fact that $\Omega \subseteq$ cl(int (Ω)). From our assumption, we know $\lambda_{\min}[\nabla^2 f(x)] < 0$ and $\exists d \in \mathbb{R}^n, d^T\nabla^2 f(x)d < 0$. By continuity, $\exists \epsilon > 0$ such that $d^T\nabla^2 f(y)d < 0, \forall y \in \bar{B}(x, \epsilon)$. Take $\tilde{x} = x + \epsilon d$. Then,

$$f(\tilde{x}) = f(x) + \nabla f(x)^T(\tilde{x} - x) + \frac{1}{2}(\tilde{x} - x)^T\nabla^2 f(y)(\tilde{x} - x)$$

for $y = x + t(\tilde{x} - x) \in \bar{B}(x, \epsilon)$ and $t \in (0, 1)$ and hence

$$f(\tilde{x}) < f(x) + \nabla f(x)^T(\tilde{x} - x)$$

(c) Same as (a) except we use strictness. □

Corollary 1.3. *Assume $\Omega \subseteq \mathbb{R}^n$ is convex, $f \in \mathcal{C}^2(\Omega)$. For $m, M \in \mathbb{R}$, we have*

$$\begin{aligned}
&mI \leq \nabla^2 f(x) \leq MI \\
\iff &f(\cdot) - \frac{m}{2}\|\cdot\|^2 \text{ and } \frac{M}{2}\|\cdot\|^2 - f(\cdot) \text{ are convex} \\
\iff &\frac{m}{2}\|y-x\|^2 \leq f(y) - [f(x) + \nabla f(x)^T(y-x)] \leq \frac{M}{2}\|y-x\|^2 \\
\iff &\frac{m}{2}\|y-x\|^2 \leq [\nabla f(y) - \nabla f(x)]^T(y-x) \leq \frac{M}{2}\|y-x\|^2
\end{aligned}$$

2 Algorithms

Definition 2.1. $d \in \mathbb{R}^n$ is a **descent direction** at x if $\exists \delta > 0$ such that $\forall t \in (0, \delta)$ we have $f(x + td) < f(x)$.

Lemma 2.1. *If $\nabla f(x)^T d < 0$ then d is a descent direction at x.*

Example 2.1. We may select $d = -\nabla f(x)$ or $d = -D\nabla f(x)$ where $D \succ 0$ as long as $\nabla f(x) \neq 0$.

Definition 2.2. A **line search method** is an algorithm with an update of the form

$$x_{k+1} = x_k + \alpha_k d_k$$

where d_k is a descent direction at x_k and α_k is a positive step size.

Definition 2.3. The **trust region method** has the following principle:

$$\alpha_k \overset{?}{=} \underset{t \in [0,\bar{\alpha}]}{\operatorname{argmin}} \{f(x_k + d_k) : t > 0\}.$$

That is, given $x_k \in \mathbb{R}^n$ we approximate $f(x_k + p) \approx m_k(p)$ where $m_k(p)$ is a simple function (e.g. $f(x_h) + \nabla f(x_h)^T p$) and solve $p_k = \operatorname{argmin}_{p \in T_k \subseteq \mathbb{R}^n}\{m_k(p)\}$ (e.g. $T_k = \bar{B}(0, \delta_k)$). If $f(x_k + p_k)$ is close to $m_k(p_k)$ then we iterate

$$x_{k+1} = x_k + p_k.$$

Otherwise, we reject $x_k + p_k$ with $x_{k+1} = x_k$ and shrink T_k. Closeness can be defined with

$$\rho_k = \frac{m_k(0) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} = \frac{f(x_k) - f(x_k + p_k)}{f(x_k) - m_k(p_k)}$$

where $\rho_k \approx 1$ implies that our estimate is close.

2.1 Steepest Descent

Definition 2.4. For a function $f \in \mathcal{C}^1(\mathbb{R}^n)$ which has $L-$Lipschitz continuous gradient, the **steepest descent with fixed step size** method is that for given $x_0 \in \mathbb{R}^n$ and $\theta \in (0, 2)$, we update with

$$\begin{aligned}
x_k &= x_{k-1} - \frac{\theta}{L}\nabla f(x_{k-1}) \\
k &\leftarrow k + 1
\end{aligned}$$

Proposition 2.1. *Assume that $f(x_k) \geq \underline{f}$ in a steepest descent method. Then for all $k > 1$ we have*

$$\min_{1 \leq i \leq k} \|\nabla f(x_{i-1})\|^2 \leq \frac{f(x_0) - \underline{f}}{k}\left(\frac{2L}{\theta(2-\theta)}\right)$$

Proof. For all $i \geq 1$, using our update step, we have

$$\begin{aligned} f(x_i) - f(x_{i-1}) &\leq \nabla f(x_{i-1})^T (x_i - x_{i-1}) + \frac{L}{2}\|x_i - x_{i-1}\|^2 \\ &\leq -\frac{\theta}{L}\|\nabla f(x_{i-1})\|^2 + \frac{\theta^2}{2L}\|\nabla f(x_{i-1})\|^2 \\ &= -\frac{\theta}{L}\|\nabla f(x_{i-1})\|^2 \left(1 - \frac{\theta}{2}\right) \end{aligned}$$

So$f(x_{i-1}) - f(x_i) \geq \frac{\theta(2-\theta)}{L}\|\nabla f(x_{i-1})\|^2$ and summing for $i = 1, 2, ..., k$ we get

$$\begin{aligned} f(x_0) - \underline{f} \geq f(x_0) - f(x_k) \geq & \frac{\theta(2-\theta)}{2L} \sum_{i=1}^{k} \|\nabla f(x_{i-1})\|^2 \\ \geq & \frac{k\theta(2-\theta)}{2L} \min_{i=1,2,\ldots,k} \|\nabla f(x_{i-1})\|^2. \end{aligned}$$

The result follows after a simple re-arrangement. □

Definition 2.5. For $\Omega \subseteq \mathbb{R}^n$ convex, $f \in \mathcal{C}^1(\Omega)$ which has $L-$Lipschitz continuous gradient on Ω, the **projected gradient** method is that for given $x_0 \in \mathbb{R}^n$ and $\theta \in (0, 2)$, we update with

$$\begin{aligned} x_k &= \operatorname*{argmin}_x \left\{ l_f(x; x_{k-1}) + \frac{L}{2\theta}\|x - x_{k-1}\|^2, x \in \Omega \right\} \\ k &\hookleftarrow k + 1 \end{aligned}$$

where $l_f(x; x_{k-1}) = f(x_{k-1}) + \nabla f(x_{k-1})^T(x - x_{k-1})$.

Lemma 2.2. *For all $k \geq 1$, under the projected gradient scheme, we have*

$$0 \in \nabla f(x_{k-1}) + N_\Omega(x_k) + \frac{L}{\theta}(x_k - x_{k-1})$$

Proof. Define $\varphi_k(x) = l_f(x; x_{k-1}) + \frac{L}{2\theta}\|x - x_{k-1}\|^2$. We first know that $\nabla_x l_f(x; x_{k-1}) = \nabla f(x_{k-1})$ and $\nabla_x \left[\frac{L}{2\theta}\|x - x_{k-1}\|^2\right] = \frac{L}{\theta}(x - x_{k-1})$ and x_k is optimal if $0 \in \nabla_x \varphi_k(x_k) + N_\Omega(x_k)$. Hence, we must have

$$0 \in \nabla f(x_{k-1}) + N_\Omega(x_k) + \frac{L}{\theta}(x_k - x_{k-1}).$$

□

Lemma 2.3. *Let $r_k = \frac{L}{\theta}(x_{k-1} - x_k)$ and $\bar{r}_k = r_k + \nabla f(x_k) - \nabla f(x_{k-1})$. Then*

$$\bar{r}_k \in \nabla f(x_k) + N_\Omega(x_k)$$

and

$$\|\bar{r}_k\| \leq L\left(\frac{1}{\theta} + 1\right)\|x_k - x_{k-1}\|$$

Proof. (Simple algebraic manipulation using Lipschitz property.) □

Remark 2.1. If we can show that $\liminf_{k\to\infty} \|\bar{r}_k\| = 0$ then the optimality condition $0 \in \nabla f(\bar{x}) + N_\Omega(\bar{x})$ is approached via $\{x_k\}$.

Lemma 2.4. *We have*

$$f(x_{k-1}) - f(x_k) \geq \frac{L}{2}\left(\frac{2-\theta}{\theta}\right)\|x_k - x_{k-1}\|^2$$

Proof. (will be shown next class) □

Proposition 2.2. *Assume that $f(x_k) \geq \underline{f}$ for all $k \geq 0$. Then, for all $k \geq 1$ we have*

$$\min_{1 \leq i \leq k} \|\bar{r}_i\|^2 \leq \frac{f(x_0) - \underline{f}}{k} \left(\frac{2L(\theta+1)^2}{\theta(2-\theta)} \right)$$

Proof. We have

$$\begin{aligned} f(x_0) - \underline{f} &\geq f(x_0) - f(x_k) \\ &= \sum_{i=1}^{k} \left(f(x_{i-1}) - f(x_i) \right) \\ &\geq \frac{L}{2} \left(\frac{2-\theta}{\theta} \right) \sum_{i=1}^{k} \|x_i - x_{i-1}\|^2 \\ &\geq \frac{L}{2} \left(\frac{2-\theta}{\theta} \right) k \min_{1 \leq i \leq k} \|x_i - x_{i-1}\|^2 \\ &\geq \frac{L}{2} \left(\frac{2-\theta}{\theta} \right) k L^2 \left(\frac{1}{\theta} + 1 \right)^2 \min_{1 \leq i \leq k} \|\bar{r}_i\|^2 \end{aligned}$$

□

2.2 Projected Gradient Method

Definition 2.6. For a space $\Omega \subseteq \mathbb{R}^n$ which is closed and convex, a function $f \in \mathcal{C}^1(\Omega)$, which has $L-$Lipschitz continuous gradient, the **linear approximation of** f is defined as

$$l_f(\tilde{x}; x) := f(x) + \nabla f(x)^T (\tilde{x} - x)$$

where $\nabla l_f(\tilde{x}; x) = \nabla f(x), l_f(x; x) = f(x)$. We have previously seen

$$|f(\tilde{x}) - l_f(\tilde{x}; x)| \leq \frac{L}{2} \|\tilde{x} - x\|^2, \forall x, \tilde{x} \in \Omega.$$

Definition 2.7. Given a space $\Omega \subseteq \mathbb{R}^n$ which is closed and convex, a function $f \in \mathcal{C}^1(\Omega)$, which has $L-$Lipschitz continuous gradient, a point $x_0 \in \Omega$, and $\theta \in (0, 2)$, the **projected gradient method** is

$$\begin{aligned} x_k &= \operatorname*{argmin}_x \left\{ l_f(x; x_{k-1}) + \frac{L}{2\theta} \|x - x_{k-1}\|^2 \right\} \qquad (1) \\ k &\hookleftarrow k + 1 \end{aligned}$$

Lemma 2.5. *For all $k \geq 1$, we have*

$$f(x_k) - f(x_{k-1}) \geq \frac{L}{2} \left(\frac{2-\theta}{\theta} \right) \|x_k - x_{k-1}\|^2$$

Proof. By (1),

$$l_f(x; x_{k-1}) + \frac{L}{2\theta} \|x - x_{k-1}\|^2 \geq l_f(x_k; x_{k-1}) + \frac{L}{2\theta} \|x_k - x_{k-1}\|^2 + \frac{L}{2\theta} \|x - x_k\|^2. \qquad (2)$$

Taking $x = x_{k-1}$,

$$\begin{aligned} f(x_{k-1}) &\geq l_f(x_k; x_{k-1}) + \frac{L}{\theta} \|x_k - x_{k-1}\|^2 \\ &= l_f(x_k; x_{k-1}) + \frac{L}{2} \|x_k - x_{k-1}\|^2 + \left(\frac{L}{\theta} - \frac{L}{2} \right) \|x_k - x_{k-1}\|^2 \\ &\geq f(x_k) + \frac{L}{2} \left(\frac{2-\theta}{\theta} \right) \|x_k - x_{k-1}\|^2 \end{aligned}$$

□

Lemma 2.6. *Given a space $\Omega \subseteq \mathbb{R}^n$ which is closed and convex, a convex function $f \in \mathcal{C}^1(\Omega)$, which has $L-$Lipschitz continuous gradient, and the set of optimal solutions $\Omega^* \neq \emptyset$ for the optimization problem*

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x \in \Omega, \end{aligned}$$

for every $k \geq 1$ and $x^* \in \Omega^*$ we have

$$\frac{L}{2}\left(\|x^* - x_{k-1}\|^2 - \|x^* - x_k\|^2\right) \geq f(x_k) - f^*.$$

Proof. By (2), with $\theta = 1$ and $x = x^*$, we have

$$\underbrace{l_f(x^*; x_{k-1}) + \frac{L}{2}\|x^* - x_{k-1}\|^2}_{\leq f(x^*) + \frac{L}{2}\|x^* - x_{k-1}\|^2} \geq \underbrace{l_f(x_k; x_{k-1}) + \frac{L}{2}\|x_k - x_{k-1}\|^2}_{\geq f(x_k) + \frac{L}{2}\|x^* - x_k\|} + \frac{L}{2}\|x^* - x_k\|^2 \tag{2}$$

and the result follows after an algebraic re-arrangement. □

Lemma 2.7. *Under the previous lemma's assumptions, for all $k \geq 1$ and $x^* \in \Omega^*$, we have*

$$\frac{L}{2}\left(\|x^* - x_0\|^2 - \|x^* - x_k\|^2\right) \geq \sum_{i=1}^{k} [f(x_i) - f^*] \geq k \cdot [f(x_k) - f^*]$$

Proof. (easy exercise) □

Lemma 2.8. *Under the previous lemma's assumptions, for all $k \geq 1$ and $x^* \in \Omega^*$, we have*

$$\begin{aligned} \|x_k - x^*\| &\leq \|x_0 - x^*\| \\ f(x_k) - f_* &\leq \frac{L}{2k}\|x_0 - x^*\|^2. \end{aligned}$$

Note that if $x^* = P_{\Omega^*}(x_0)$ then $d_0 := \|x_0 - P_\Omega(x^*)\|$ can be thought of a distance of x_0 to Ω^* and

$$f(x_k) - f_* \leq \frac{Ld_0^2}{2k}.$$

Proof. (follows from the previous lemma) □

Lemma 2.9. *Define*

$$\tilde{r}_k = \frac{L}{2\theta}(x_{k-1} - x_k) + \nabla f(x_k) - \nabla f(x_{k-1}).$$

Then $r_k \in \nabla f(x_k) + N_\Omega(X_k)$ where if $r_k = 0$ then x_k satisfies the optimality condition of

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

Proof. Left as an exercise (?) □

Definition 2.8. $\{a_k\}_{k=1}^\infty \subseteq \mathbb{R}$ **converges geometrically** if there exists $\gamma \geq 0$ and $\tau \in (0, 1)$ such that

$$a_k \leq \gamma\tau^k, \forall k \geq 1.$$

Note 1. $\lim_{k\to\infty}(a_k/[1/k^p]) = 0$ for $p > 0$, but the rate at which a_k diminishes may be REALLY slow relative to $1/k^p$.

Lemma 2.10. *Given a space $\Omega \subseteq \mathbb{R}^n$ which is closed and convex, a $\beta-$strongly convex function $f \in \mathcal{C}^1(\Omega)$, which has $L-$Lipschitz continuous gradient, and the set of optimal solutions $\Omega^* \neq \emptyset$ for the optimization problem*

$$\begin{aligned} &\min\, f(x) \\ &\textit{s.t. } x \in \Omega, \end{aligned}$$

for every $k \geq 1$ and $x^* \in \Omega^*$ we have

$$\frac{L}{2}\left(1 - \frac{\beta}{2}\right)^k d_0^2 \geq f(x_k) - f^*.$$

Proof. By (2), with $\theta = 1$ and $x = x^*$, we have

$$\underbrace{l_f(x^*; x_{k-1}) + \frac{L}{2}\|x^* - x_{k-1}\|^2}_{\leq f(x^*) + \frac{(1-\beta)}{2}\|x^* - x_{k-1}\|^2} \geq \underbrace{l_f(x_k; x_{k-1}) + \frac{L}{2}\|x_k - x_{k-1}\|^2}_{\geq f(x_k) + \frac{L}{2}\|x^* - x_k\|} + \frac{L}{2}\|x^* - x_k\|^2 \tag{2}$$

and the result follows after an algebraic re-arrangement and iterating over k. □

Exercise 2.1. Recall $r_k \in \nabla f(x_k) + N_\Omega(X_k)$ and

$$\|\tilde{r}_k\| \leq L\left(1 + \frac{1}{\theta}\right)\|x_k - x_{k-1}\|.$$

For $\theta = 1$, we have

$$\|\tilde{r}_k\| \leq 2L\|x_k - x_{k-1}\|.$$

Show that

$$\begin{aligned} \min_{i=1,\ldots,k} \|\tilde{r}_i\|^2 &= \mathcal{O}\left(\frac{1}{k^2}\right) \text{ if } f \text{ is convex} \\ \|\tilde{r}_k\| &= \mathcal{O}\left(\left(1 - \frac{\beta}{L}\right)^k\right) \text{ if } f \text{ is } \beta-\text{strongly convex} \end{aligned}$$

Remark 2.2. For a function $f(x) = \frac{1}{2}(x - x^*)^T Q(x - x^*) + \gamma$, we have $L = \lambda_{\max}(Q)$, $\beta = \lambda_{\min}(Q)$ and $\text{cond}(Q) = \lambda_{\max}(Q)/\lambda_{\min}(Q)$ so $\|\tilde{r}_k\|$ is related to the inverse condition number of Q.

2.3 Gradient-type Methods

Problem 2.1. For standard minimization algorithms of the form $x_{k+1} = x_k + \alpha_k d_k$ where α_k, d_k are respective step sizes and descent directions, what conditions on $\{\alpha_k\}, \{d_k\}$ should we set to ensure convergence?

Remark 2.3. Assuming the function is still $L-$Lipschitz, we know:

$$\begin{aligned} f(x') - f(x) - \nabla f(x)^T(x' - x) &\leq \frac{L}{2}\|x' - x\|^2 \\ \implies f(x_k + \alpha d_k) - f(x_k) &\leq \alpha \nabla f(x_k)^T d_k + \frac{L}{2}\|d_k\|^2. \end{aligned}$$

Take

$$\alpha_k = \operatorname*{argmin}_{\alpha \in \mathbb{R}} \left\{\alpha \nabla f(x_k)^T d_k + \frac{L\alpha^2}{2}\|d_k\|^2\right\}$$

where at optimality, we need

$$\begin{aligned} \nabla f(x_k)^T d_k + \alpha_k\|d_k\|^2 &= 0 \\ \implies \alpha_k &= -\frac{\nabla f(x_k)^T d_k}{L\|d_k\|^2} > 0. \end{aligned}$$

Substituting this into the Lipschitz condition yields

$$f(x_k) - f(x_{k+1}) \geq \frac{(\nabla f(x_k)^T d_k)^2}{2L\|d_k\|^2} > 0.$$

Remark 2.4. Let $\epsilon_k = \frac{-\nabla f(x_k)^T d_k}{\|\nabla f(x_k)\|\|d_k\|}$ where $\epsilon_k = \cos\theta_k$ and θ_k is the angle between d_k and $-\nabla f(x_k)$. Then,

$$\begin{aligned}
f(x_k) - f(x_{k+1}) &\geq \frac{\epsilon_k^2 \|\nabla f(x_k)\|^2}{2L} \\
\implies f(x_0) - \underline{f} \geq f(x_0) - f(x_k) &\geq \sum_{i=0}^{k-1} f(x_i) - f(x_{i+1}) \geq \sum_{i=0}^{k-1} \frac{\epsilon_i^2 \|\nabla f(x_i)\|^2}{2L} \\
\implies f(x_0) - \underline{f} &\geq \frac{1}{2L}\left(\min_{i\leq k-1} \|\nabla f(x_i)\|^2\right)\left(\sum_{i=0}^{k-1} \epsilon_i^2\right) \\
\implies \min_{i\leq k-1} \|\nabla f(x_i)\|^2 &\leq \frac{2L\left(f(x_0) - \underline{f}\right)}{\sum_{i=0}^{k-1} \epsilon_i^2}.
\end{aligned}$$

So if $\sum_{i=0}^{\infty} \epsilon_i^2 = \infty$ (e.g. $\epsilon_i \geq \underline{\epsilon}$ for all i), then $\lim_{k\to\infty} \min_{i\leq k} \|\nabla f(x_i)\|^2 = 0$ or $\liminf_{h\to\infty} \|\nabla f(x_k)\| = 0$. If $\epsilon_i \geq \epsilon$ for all i, then

$$\min_{i\leq k-1} \|\nabla f(x_i)\|^2 \leq \frac{2L\left(f(x_0) - \underline{f}\right)}{\epsilon^2 k}.$$

Exercise 2.2. If $\alpha_k = -\theta \frac{\nabla f(x_k)^T d_k}{L\|d_k\|^2}$ and $\theta \in (0, 2)$, show that

$$f(x_k) - f(x_{k+1}) \geq \left(\theta - \frac{\theta^2}{2}\right)\left(\frac{(\nabla f(x_k)^T d_k)^2}{L\|d_k\|^2}\right).$$

Remark 2.5. If $d_k = -D_k \nabla f(x_k)$ and D_k is symmetric positive definite, then $\text{cond}(D_n) \leq \frac{1}{\epsilon} \implies \epsilon_k \geq \epsilon > 0$. The proof makes use of the fact that

$$\begin{gathered}
\lambda_{\min}(D)\|u\|^2 \leq u^T D u \leq \lambda_{\max}(D)\|u\|^2 \\
\|Du\| \leq \lambda_{\max}(D)\|u\|
\end{gathered}$$

and with $g = \nabla f(x)$, we have

$$\epsilon_k = -\frac{g_k^T d_k}{\|g_k\|\|d_k\|} = \frac{g_k^T D_k g_k}{\|g_k\|\|d_k\|} \geq \frac{\lambda_{\min}(D_k)\|g_k\|^2}{\|g_k\|\lambda_{\max}(D_k)\|g_k\|} = \frac{1}{\text{cond}(D_k)} \geq \epsilon.$$

2.4 Inexact Line Search

Remark 2.6. Assume now that L is not known or does not exist and define $\phi_k(\alpha) = f(x_k + \alpha d_k) - f(x_k)$. We wish to choose α such that

$$\phi_k(\alpha) \leq \sigma \phi_k'(0) \cdot \alpha$$

where $\sigma \in (0, 1)$ is a fixed constant, where we wish "to not be close to $\bar{\alpha}$, a root of ϕ". To not be close to 0, there are many strategies:

- (a) Goldstein rule: For some constant $\tau \in (\sigma, 1)$, we require α_k to satisfy

$$\phi_k(\alpha) \geq \tau \phi_k'(0)\alpha \qquad (*)$$

- (b) Wolfe-Powell (W-P) rule: For some constant $\tau \in (\sigma, 1)$, we require α_k to satisfy

$$\phi_k'(\alpha) \geq \tau \phi_k'(0)$$

- (c) Strong Wolfe-Powell rule: For some constant $\tau \in (\sigma, 1)$, we require α_k to satisfy

$$|\phi_k'(\alpha)| \leq -\tau\phi_k'(0)$$

with $\sigma < \frac{1}{2}$.

- (d) Armijo's rule: Let $s > 0$ and $\beta \in (0, 1)$ be fixed constants. Choose α_k as the largest scalar from

$$\alpha \in \{s, s\beta, s\beta^2, ...\}$$

such that (*) is satisfied.

Proposition 2.3. *With respect to Armijo's rule,*

1) $\exists\delta > 0$ such that (*) is satisfied strictly for any $\alpha \in (0, \delta)$.

2) If $\{\phi_k(\alpha) : \alpha > 0\}$ is bounded below, there exists an open interval of α's that satisfy rules (a) to (c).

Proof. Left as an exercise. □

Theorem 2.1. *Suppose that*

1) $f \in C^1(\mathbb{R}^n)$ and there exists $L > 0$ such that for all $y, z \in \{x : f(x) \leq f(x^0)\}$ we have

$$\|\nabla f(y) - \nabla f(x)\| \leq L\|y - x\|.$$

2) $\{f(x_k)\}$ is bounded from below.

3) $\{d_k\}$ is gradient-related if α_k is chosen by Armijo's rule, i.e., there exists $\delta > 0$ such that

$$\|d_k\| \geq \delta\|\nabla f(x_k)\|, \forall k \geq 0$$

Then,

$$\sum_{k=0}^{\infty} \epsilon_k^2 \|\nabla f(x_k)\|^2 < \infty$$

and hence if $\sum_{i=0}^{\infty} \epsilon_i^2 = \infty$ then

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

Thus, every accumulation point of $\{x_k\}$ is a stationary point.

Rates of Convergence

Consider the problem $\min_{x\in\mathbb{R}^n} \left\{f(x) = \frac{1}{2}x^TQx + c^Tx + \gamma\right\}$ where $Q > 0$ is symmetric.

Steepest Descent

The algorithm for our problem is

$$\begin{aligned} x_{k+1} &= x_k - \alpha_k g_k \\ g_k &= \nabla f(x_k) \\ \alpha_k &= \operatorname*{argmin}_{\alpha\in\mathbb{R}} f(x^k - \alpha g_k) = \frac{\|g_k\|^2}{g_k^TQg_k} \end{aligned}$$

Proposition 2.4. *For every $k \geq 0$, we have*

$$\frac{f(x_{k+1}) - f_*}{f(x_k) - f_*} \leq \left(\frac{M - m}{M + m}\right)^2 = \left(\frac{r - 1}{r + 1}\right)^2$$

where $m = \lambda_{\min}(Q), M = \lambda_{\max}(Q)$ and $r = M/m =$ cond$(Q) \geq 1$.

Proof. (see related proof for the projected gradient)

Gradient-type Methods

The algorithm for our problem is

$$
\begin{aligned}
x_{k+1} &= x_k - \alpha_k D_k g_k \text{ where } D_k > 0 \\
\alpha_k &= \operatorname*{argmin}_{\alpha \in \mathbb{R}} f(x_k - \alpha D_k g_k)
\end{aligned}
$$

□

Proposition 2.5. *For every $k \geq 0$, we have*

$$
\frac{f(x_{k+1}) - f_*}{f(x_k) - f_*} \leq \left(\frac{M_k - m_k}{M_k - m_k} \right)^2 = \left(\frac{r_k - 1}{r_k + 1} \right)^2
$$

where $M_k = \lambda_{\max}(D_k^{1/2} Q D_k^{1/2}), m_k = \lambda_{\max}(D_k^{1/2} Q D_k^{1/2})$, and $r_k =$ cond$(D_k^{1/2} Q D_k^{1/2})$.

Proof. We first note that

$$
\begin{aligned}
0 = \frac{d}{d\alpha} f(x_k + \alpha d_k) = \nabla f(x_k + \alpha d_k)^T d_k &= [\nabla f(x_k) + \alpha_k Q d_k]^T d_k \\
&= \nabla f(x_k^T) d_k + \alpha_k d_k^T Q d_k
\end{aligned}
$$

implies that $\alpha_k = -\frac{\nabla f(x_k)^T d_k}{d_k^T Q d_k}$. Next, if we define $\tilde{f}(y) = f(Sy)$ where $s = D_k^{1/2}$ then

$$
\begin{aligned}
\nabla \tilde{f}(y) &= S \nabla f(Sy) \\
\nabla^2 f(y) &= S \nabla^2 f(Sy) S = SQS.
\end{aligned}
$$

For every k let $y = S^{-1} x_k$ and note by our iteration scheme we have $\nabla \tilde{f}(y_k) = S \nabla f(x_k) = S g_k$ as well as

$$
S y_{k+1} = S y_k - \alpha_k S^2 \nabla f(S y_k) \implies y_{k+1} = y_k - \alpha_k \nabla \tilde{f}(y_k)
$$

and

$$
\begin{aligned}
\alpha_k &= \operatorname*{argmin}_{\alpha \in \mathbb{R}} f(x_k - \alpha D_k g_k) \\
&= \operatorname*{argmin}_{\alpha \in \mathbb{R}} \tilde{f}(y_k - \alpha S g_k) \\
&= \operatorname*{argmin}_{\alpha \in \mathbb{R}} \tilde{f}(y_k - \alpha \nabla \tilde{f}(y_k)).
\end{aligned}
$$

From the previous proposition,

$$
\frac{\tilde{f}(x_{k+1}) - f_*}{\tilde{f}(x_k) - f_*} \leq \left(\frac{M_k - m_k}{M_k - m_k} \right)^2 = \left(\frac{r_k - 1}{r_k + 1} \right)^2
$$

where $M_k = \lambda_{\max}(D_k^{1/2} Q D_k^{1/2}), m_k = \lambda_{\max}(D_k^{1/2} Q D_k^{1/2})$, and $r_k = \text{cond}(D_k^{1/2} Q D_k^{1/2})$. □

Remark 2.7. If $r_k \to 1$ then

$$
\lim_{k \to \infty} \frac{f(x_{k+1}) - f_*}{f(x_k) - f_*} = 0.
$$

For example, if $D_k \to Q^{-1}$, then the above holds.

2.5 Newton's Method

Newton's Method

Consider a function $h : \mathbb{R}^n \mapsto \mathbb{R}^n$ where $h \in \mathcal{C}^1(\mathbb{R}^n)$. Newton's method finds a point $x \in \mathbb{R}^n$ where $h(x) = 0$. The idea for a given x_k, uses

$$h(x) \approx h(x_k) + h'(x_k)(x - x_k) = 0 \implies x_{k+1} = x_k - h'(x_k)^{-1}h(x_k).$$

In the case of $h(x) = \nabla f(x) = 0$ where $h'(x) = \nabla^2 f(x)$, we have the iteration scheme

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

In general optimization, we may use a second order approximation to $f(x)$ and apply Newton's method to find where $\nabla f(x) = 0$.

Local Convergence of Newton's Method

Theorem 2.2. *Assume $h \in \mathcal{C}^2(\mathbb{R}^n)$ and let $x^* \in \mathbb{R}^n$ be such that $h(x^*) = 0, h'(x^*)$ is non-singular. Then there exists $y > 0$ such that if $x_0 \in \bar{B}(x^*; y)$ then $\{x_k\}$ obtained as*

$$x_{k+1} = x_k - [h'(x_k)]^{-1}h(x_k)$$

is well-defined and

$$\lim_{k\to\infty} x_k = x^* \textit{ and } \limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} < \infty.$$

Proof. Let $L := 2\|h'(x^*)^{-1}\|$ and choose $y > 0$ such that for all $x \in \bar{B}(x^*; y)$ we have

- $h'(x)$ exists, $\|h'(x)^{-1}\| \leq L$
- $\frac{\eta LM}{2} < 1$ where $M = \sup_{x\in\bar{B}(x^*;y)} \|h''(x)\|$

It can be shown that

$$\|h'(x) - h'(y)\| \leq M\|x - y\|, \forall x, y \in \bar{B}(x^*; y).$$

Then if $x_k \in \bar{B}(x^*; y)$ we have

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - h'(x_k)^{-1}h(x_k) \\ &= h'(x_k)^{-1}\left[\underbrace{h(x^*)}_{=0} - h(x_k) - h'(x_k)^{-1}h(x_k)\right]. \end{aligned}$$

So

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \|h'(x_k)^{-1}\|\|h(x^*) - h(x_k) - h'(x_k)(x^* - x_k)\| \\ &\leq L\left\|\int_0^1 [h'(x_k + t(x^* - x_k)) - h'(x_k)]\,(x^* - x_k)\,dt\right\| \\ &\leq L\int_0^1 \|h'(x_k + t(x^* - x_k)) - h'(x_k)\|\,\|x^* - x_k\|\,dt \\ &\leq L\|x^* - x_k\|\int_0^1 Mt\|x^* - x_k\|\,dt \\ &= \frac{ML}{2}\|x^* - x_k\|^2 \\ &\leq \frac{ML\mu}{2}\|x^* - x_k\| \\ &< \|x_k - x^*\| \end{aligned}$$

and hence

$$\lim_{k\to\infty} \|x_k - x^*\| = 0.$$

□

2.6 Conjugate Gradient Method

Suppose we are dealing with the problem $\min_{x\in\mathbb{R}^n}\left\{\frac{1}{2}x^TQx - b^Tx\right\}$ where $Q > 0$ is symmetric. Let $\{d_0, ..., d_n\}$ be a basis for $\mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, and denote $[d_0, ..., d_k]$ as the subspace spanned by $d_0, ..., d_k$. **We use the notation** $g_k = \nabla f(x_k)$.

Lemma 2.11. *We have*

$$x_{k+1} = \operatorname{argmin}\left\{f(x) = \frac{1}{2}x^TQx - b^Tx : x \in x_0 + [d_1, ..., d_k]\right\}$$

if and only if

$$x_{k+1} = x_0 - D_k\left(D_k^TQD_k\right)^{-1}D_k^Tg_0$$

where

$$D_k = [d_0 ... d_k] \in \mathbb{R}^{n\times(k+1)}$$
$$g_0 = \nabla f(x_0) = Qx_0 - b.$$

Also, $g_{k+1} \perp d_i$ *for* $i = 0, ..., k$.

Proof. We know that

$$x \in x_0 + [d_0, ..., d_k] \iff x = x_0 + D_kz, \text{ for some } z \in \mathbb{R}^{k+1}$$

So $x_{k+1} = x_0 + D_kz_{k+1}$ where $z_{k+1} = \operatorname{argmin}_z f(x_0 + D_kz) = h(z)$. In particular, z_{k+1} solves

$$\begin{aligned} 0 = \nabla h(z) &= D_k^T\nabla f(x_0 + D_kjz) \\ &= D_k^T\left[Q\left(x_0 + D_kz\right) - b\right] \\ &= (D_k^TQD_k)z + D_k^Tg_0. \end{aligned}$$

So, $z_{k+1} = -(D_k^TQD_k)^{-1}D_k^Tg_0$ and the result follows after re-arranging terms and remarking that

$$0 = D_k^T\nabla f(x_0 + D_kz_{k+1}) = D_k^Tg_{k+1}.$$

□

Definition 2.9. A set of directions $\{d_0, ..., d_k\} \subseteq \mathbb{R}^n$ are Q**-conjugate** if $d_i^TQd_j = 0$ for every $0 \le i < j \le k$. Equivalently, $D_k^TQD_k$ is diagonal.

Proposition 2.6. *Suppose that* $Q > 0$ *and* $d_0, ..., d_k$ *are* Q*-conjugate vectors. Then* $d_0, ..., d_k$ *are linearly independent.*

Proof. Exercise. □

Theorem 2.3. *(Expanding Subspace Minimization) Assume that* $x_{k+1} = \operatorname{argmin}\{f(x) : x \in x_0 + [d_0, ..., d_k]\}$ *and that* $d_0, ..., d_{k-1}$ *are* Q*-conjugate. Then,*

(a) $x_n = x^*$

(b) $g_{k+1}^Td_i = 0$ *for* $i = 0, ..., k$, $k \ge 1$

(c) $x_{k+1} = x_k + \alpha_kd_k$ *where* $\alpha_k = -\frac{g_k^Td_k}{d_k^TQd_k}$

or equivalently $\alpha_k = \operatorname{argmin} f(x_k + \alpha d_k)$

or equivalently $x_{k+1} = \operatorname{argmin}\{f(x) : x \in x_k + [d_k]\}$.

Proof. (a) and (b) are obvious. For (c), note that

$$x_k \in x_0 + [d_0, ..., d_{k-1}] \subseteq x_0 + [d_0, ..., d_k]$$

and so

$$x_k + [d_0, ..., d_k] = x_0 + [d_0, ..., d_k].$$

In the previous algorithms, we can hence replace x_0 with x_k. In particular, the first lemma can be replaced with the iteration scheme

$$x_{k+1} = x_k - D_k \left(D_k^T Q D_k\right)^{-1} D_k^T g_k.$$

Simplifying with the fact that

$$D_k^T g_k = (g_k^T d_k) e_{k+1}$$
$$D_k^T Q D_k = \text{diag}(d_1^T Q d_1, ..., d_k^T Q d_k)$$
$$(D_k^T Q D_k)^{-1} D_k^T g_k = \frac{g_k^T d_k}{d_k^T Q d_k^T} \cdot e_{k+1}$$

where e_{k+1} is the $(k+1)^{th}$ basis vector in $\mathbb{R}^n$, this then reduces the iteration scheme further to

$$x_{k+1} = x_k - \left(\frac{g_k^T d_k}{d_k^T Q d_k^T}\right) D_k e_{k+1} = x_k - \left(\frac{g_k^T d_k}{d_k^T Q d_k^T}\right) d_k = x_k - \alpha_k d_k.$$

□

Algorithm 1. *(Conjugate Gradient Method [sketch]) Given $x_0 \in \mathbb{R}^n$, let $d_0 = -g_0 = b - Qx_0$. For $k = 0, 1, 2, ...$ do*

$$x_{k+1} = x_k + \alpha_k d_k \text{ where } \alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k}.$$

If $g_{k+1} = 0$, stop; else $d_{k+1} = -g_{k+1} + \beta_k d_k$ where $\beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k}$.

Remark 2.8. Observe that

$$0 = d_{k+1}^T Q d_k = (-d_{k+1} + \beta_k d_k)^T Q d_k = -g_{k+1}^T Q d_k + \beta_k d_k^T Q d_k \implies \beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k}$$

Lemma 2.12. *(Gram-Schmidt) Assume that $d_0, ..., d_{i-1}$ are Q-conjugate nonzero vectors and $p_i \notin [d_0, ..., d_{k-1}]$. Define*

$$d_k = p_k - \sum_{i=0}^{k-1} \frac{p_k^T Q d_i}{d_i^T Q d_i} d_i = p_k + \sum_{i=0}^{k-1} \beta_{k-1,i} d_i \text{ where } \beta_{k-1} = -\frac{p_k^T Q d_i}{d_i Q d_i}.$$

Then $d_0, ..., d_k$ are Q-conjugate nonzero vectors and

$$[d_0, ..., d_k] = [d_0, ..., d_{k-1} p_k].$$

Proof. Exercise. □

Algorithm 2. *(Alternate Conjugate Gradient) For $x_0 \in \mathbb{R}^n$, $f(x) = \frac{1}{2} x^T Q x - b^T x$, $Q > 0$ symmetric, let $d_0 = -g_0 = b - Qx_0$. For $k = 0, 1, 2, ...$ do*

$$x_{k+1} = x_k + \alpha_k d_k \text{ where } \alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k}.$$

If $g_{k+1} = 0$, stop; else $d_{k+1} = -g_{k+1} + \sum_{i=1}^k \beta_{ki} d_i$ where $\beta_{ki} = \frac{g_{k+1}^T Q d_i}{d_i^T Q d_i}$. Here, we are generating the $g_k \perp [d_0, ..., d_{k-1}]$ vectors on the fly and by adapting Gram-Schmidt we have the added bonus that we are preserving Q-conjugacy.

Lemma 2.13. *If $d_0, ..., d_k$ are Q-conjugate and $g_{k+1} \notin [d_0, ..., d_k]$ then d_{k+1} as above satisfies*

(1) d_{k+1} is Q-conjugate w.r.t. $\{d_0, ..., d_k\}$

(2) $[d_0, ..., d_{k+1}] = [d_0, ..., d_k, g_{k+1}]$

Theorem 2.4. *Assume that $g_i \neq 0$, $i \in \{0, ..., h\}$. Then for all $i \in \{0, 1, ..., k\}$ we have*

(i) $d_0, ..., d_i$ are Q-conjugate

(ii) $g_0, ..., g_i$ are orthogonal

(iii) $[d_0, ..., d_i] = [g_0, ..., g_i]$

(iv) $[d_0, ..., d_i] = [g_0, Qg^0, ..., Q^i g_0]$

(v) $\alpha_i = \|g_i\|/(d_i^T Q d_i)$ *and* $g_i^T d_i = -\|g_i\|^2$

Proof. By induction on i. For $i = 0$, it is obvious. Assume it is true for $i - 1$. Hence,

(a) $[d_0, ..., d_{i-1}] = [g_0, ..., g_{i-1}] = ([g_0, Qg^0, ..., Q^{i-2} g_0] = \mathcal{L}_{i-1})$

(b) $g_0, ..., g_{i-1}$ are orthogonal

(c) $d_0, ..., d_{i-1}$ are Q-conjugate

By our previous lemma, d_i is Q-conjugate w.r.t. $\{d_0, ..., d_{i-1}\}$ and so (i) follows. Also by the lemma, we know

$$[d_0, ..., d_i] = [d_0, ..., d_{i-1}, g_i] = [g_0, ..., g_{i-1} = g_i]$$

from (a) which shows (iii).

Next, we have

$$d_i \in [d_0, ..., d_{i-1}, g_i] = [g_0, ..., Q^{i-1} g_0, g_i]$$

from (a). Also $g_i = g_{i-1} + \alpha_{i-1} Q d_{i-1}$ with $g_{i-1} \in \mathcal{L}_{i-1}$ and $Qd_{i-1} \in Q\mathcal{L}_{i-1} = \mathcal{L}_i$ so $g_i \in \mathcal{L}_i$. This tells us then that $d_i \in \mathcal{L}_i \implies [d_0, ..., d_i] \subseteq \mathcal{L}_i$. Since $d_0, ..., d_i$ are linearly independent then $[d_0, ..., d_i] = \mathcal{L}_i$ and (iv) follows.

Now we have $g_i \perp [d_0, ..., d_{i-1}]$ since the method is a Q-conjugate direction method. Since $[d_0, ..., d_{i-1}] = [g_0, ..., g_{i-1}]$ then $g_i \perp [g_0, ..., g_{i-1}]$ and (ii) follows.

For (v) note that $d_i = -g_i + u$ with $u \in \mathcal{L}_{i-1}$ and hence $g_i^T d_i = -\|g_i\|^2 + \underbrace{u^T g_i}_{=0}$ and the definition of α_i follows. □

Proposition 2.7. *Assume that* $g_{k+1} \neq 0$. *Then*

$$\beta_{ki} = \begin{cases} \frac{\|g_{k+1}\|^2}{\|g_k\|^2} & i = k \\ 0 & i < k. \end{cases}$$

Proof. By definition $\beta_{ki} = \frac{g_{k+1}^T Q d_i}{d_i^T Q d_i}$ and

$$Qd_i = Q\left(\frac{x_{i+1} - x_i}{\alpha_i}\right) = \frac{g_{i+1} - g_i}{\alpha_i} \implies g_{k+1}^T Q d_i = g_{k+1}\left(\frac{g_{i+1} - g_i}{\alpha_i}\right) = \begin{cases} \frac{\|g_{k+1}\|^2}{\alpha_k} & i = k \\ 0 & i < k. \end{cases}$$

Next,

$$d_i^T Q d_i = d_i^T \left(\frac{g_{i+1} - g_i}{\alpha_i}\right) = -\frac{d_i^T g_i}{\alpha_i} = \frac{\|g_i\|^2}{\alpha_i}$$

and the result follows.

<u>Convergence Rate of the Conjugate Gradient Method</u>

Note that

$$\begin{aligned} x \in x_0 + [d_0, ..., d_{k-1}] &\iff x \in x_0 + [g_0, ..., Q^{k-1} g_0] \\ &\iff x = x_0 + \gamma_1 g_0 + ... + \gamma_k Q^{k-1} g_0 \text{ for some } \gamma \in \mathbb{R}^k \end{aligned}$$

Now, we have $g_0 = Q(x_0 - x^*)$ and hence

$$\begin{aligned} x - x^* &= x_0 - x^* + \gamma_1 Q(x_0 - x^*) + ... + \gamma_k Q^k (x_0 - x^*) \\ &= (I + \gamma_1 Q + ... + \gamma_k Q^k)(x_0 - x^*) \\ &= P_k(Q)(x_0 - x^*) \end{aligned}$$

where $P_k \in \mathcal{P}_k$ and $\mathcal{P}_k$ is the set of polynomials of degree at most k such that $P_k(0) = 1$. Now we have $f(x) = f(x^*) + \frac{1}{2}(x - x^*)Q(x - x^*)$ so

$$f(x) - f(x^*) = \frac{1}{2}\|Q^{1/2}(x - x^*)\|^2$$

and the original QP is equivalent to

$$
\begin{aligned}
2(f(x_k) - f(x^*)) = \begin{array}{l} \min \|Q^{1/2}(x - x^*)\|^2 \\ \text{s.t. } x \in x_0 + [d_0, ..., d_{k-1}] \end{array} &= \begin{array}{l} \min \|Q^{1/2}(x - x^*)\|^2 \\ \text{s.t. } x - x^* = P_k(Q)(x_0 - x^*) \\ \quad P_k \in \mathcal{P}_k \end{array} \\
&= \begin{array}{l} \min \|Q^{1/2} P_k(Q)(x_0 - x^*)\|^2 \\ \text{s.t. } P_k \in \mathcal{P}_k \end{array} \\
&= \begin{array}{l} \min \|P_k(Q) Q^{1/2}(x_0 - x^*)\|^2 \\ \text{s.t. } P_k \in \mathcal{P}_k \end{array} \\
&\leq \left(\begin{array}{c} \min \|P_k(Q)\| \\ \text{s.t. } P_k \in \mathcal{P}_k \end{array} \right)^2 \|Q^{1/2}(x_0 - x^*)\|
\end{aligned}
$$

□

Proposition 2.8. *For every $k \geq 0$, we have*

$$\frac{f(x_k) - f_*}{f(x_0) - f_*} \leq \left(\begin{array}{c} \min \|P_k(Q)\| \\ s.t. \ P_k \in \mathcal{P}_k \end{array} \right)^2$$

and since

$$\|P_k(Q)\| = \max_{\lambda \in \sigma(Q)} |P_k(\lambda)|$$

*where $\sigma(Q)$ is the **spectrum** of Q or the set of eigenvalues of Q.*

Corollary 2.1. *For every $k \geq 0$ and $P_k \in \mathcal{P}_k$, we have*

$$\frac{f(x_k) - f_*}{f(x_0) - f_*} \leq \left(\max_{\lambda \in \sigma(Q)} |P_k(\lambda)| \right)^2.$$

Corollary 2.2. *Assume that Q has $m < n$ distinct eigenvalues. Then $x_m = x^*$.*

Proof. Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of Q. Let $P_m \in \mathcal{P}_m$ be defined as

$$P_m(\lambda) = \frac{\prod_{i=1}^m (\lambda_i - \lambda)}{\prod_{i=1}^m \lambda_i}$$

and since $P_m(\lambda) = 0$ for every $\lambda \in \sigma(Q)$ then from the previous proposition, the result follows. □

<u>Rate of Convergence of CG Method</u>

Corollary 2.3. *Assume that Q has*

(1) $(n - m)$ eigenvalues in $[a, b]$, $m > 0$

(2) m eigenvalues which are greater than b.

Then,

$$\frac{f(x_{m+1}) - f_*}{f(x_0) - f_*} \leq \left(\frac{b - a}{a + b} \right)^2.$$

In particular, for $m = 0$ and $a = \lambda_{\min}, b = \lambda_{\max}$, we have

$$\frac{f(x_1) - f_*}{f(x_0) - f_*} \leq \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right).$$

Proof. Let $\lambda_1, ..., \lambda_m$ denote the eigenvalues greater than b and define $\lambda_{m+1} = \frac{b+a}{2}$. Next, define

$$P_{m+1}(\lambda) = \frac{\prod_{i=1}^{m+1} (\lambda_i - \lambda)}{\prod_{i=1}^{m+1} \lambda_i}$$

where clearly $P_{m+1} \in \mathcal{P}_{m+1}$. By a previous proposition,

$$\frac{f(x_{m+1}) - f_*}{f(x_0) - f_*} \le \max_{\lambda \in [a,b]} |P_{m+1}(\lambda)|^2 \le \max_{\lambda \in [a,b]} \left| 1 - \frac{2\lambda}{a+b} \right|^2 = \left(\frac{b-a}{a+b} \right)^2.$$

□

Corollary 2.4. *For all $k \ge 0$, we have*

$$\frac{f(x_k) - f_*}{f(x_0) - f_*} \le 2 \left(\frac{\sqrt{r} - 1}{\sqrt{r} + 1} \right)^2$$

where $r = M/m$ is the condition number of Q.

Proof. (sketch) Use the polynomials

$$T_k(x) = \frac{1}{2} \left(x + \sqrt{x^2 - 1} \right)^k + \frac{1}{2} \left(x - \sqrt{x^2 - 1} \right)^k = \cos(k \arccos x), |T_k(x)| \le 1, \forall x \in [-1, 1]$$

and define

$$P_k(\lambda) = \frac{T_k \left(\frac{2\lambda - (m+M)}{M-m} \right)}{T_k \left(-\frac{M+m}{M-m} \right)} \in \mathcal{P}_k.$$

Use a similar procedure as before to obtain the result. □

2.7 General Conjugate Gradient Method

Definition 2.10. Consider the problem $(*)\ \min\{f(x) : x \in \mathbb{R}^n\}$ where $f \in \mathcal{C}^1(\mathbb{R}^n)$. The **CG framework**, given $x_0 \in \mathbb{R}^n$, is: For $k = 0, 1, ...$ do

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ d_{k+1} &= -\nabla f(x_{k+1}) + \beta_k d_k \end{aligned}$$

where $\alpha_k > 0$ is the step size (e.g. use an exact or inexact line search method). Recall for convex quadratic,

$$\beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k} = \underbrace{\frac{\|g_{k+1}\|^2}{\|g_k\|^2}}_{(1)} = \underbrace{\frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|^2}}_{(2)}.$$

Using (1) in the general case leads to the **Fletcher-Reeves** (FR) method while (2) leads to the **Polak-Ribière** (PR) method. Note that using (2) implies that

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 < 0.$$

Theorem 2.5. *(PR) Assume that f is such that for $0 < m \le M$,*

$$m\|u\|^2 \le u^T \nabla^2 f(x) u \le M\|u\|^2$$

for all $x, u \in \mathbb{R}^n$. Then the PR-CG method with exact line search method converges to the unique global minimum of $(*)$.

Theorem 2.6. *Assume that $f \in \mathcal{C}^2(\mathbb{R}^n)$ and $\{x : f(x) \le f(x_0)\}$ is bounded. Then there exists an accumulation point $\bar{x}$ of $\{x_k\}$ such that $\nabla f(\bar{x}) = 0$. If f is convex then $\{\bar{x}_k\} \to \bar{x}$.*

*The **Strong Wolfe-Powell inexact line search** is used in this scheme where $0 < \sigma < \frac{1}{2}$, $\sigma < \tau < 1$ and*

$$\begin{aligned} f(x_k + \alpha_k d_k) - f(x_k) &\le \sigma \alpha_k \nabla f(x_k)^T d_k \\ |\nabla f(x_k + \alpha_k d_k)^T d_k| &\le -\tau \nabla f(x_k)^T d_k \end{aligned}$$

2.8 Nesterov's Method

Definition 2.11. Suppose that $f \in \mathcal{C}^1(\mathbb{R}^n)$ is convex and $\nabla f(x)$ is L-Lipschitz where

$$l_f(\tilde{x};x) \leq f(\tilde{x}) \leq l_f(\tilde{x};x) + \frac{L}{2}\|\tilde{x}-x\|^2, l_f(\tilde{x};x) = f(x) + \nabla f(x)^T(\tilde{x}-x).$$

For the problem $\min\{f(x) : x \in X\}$, let $X^* \neq \emptyset$ be a closed and convex set of optimal set of solutions. The **Nesterov Method** is as follows:

(0) Let $x_0 \in \mathbb{R}^n$ be given and set $y_0 = x_0, k = 0, A_0 = 0$.

(1) Compute

$$\begin{aligned}
a_k &= \frac{1+\sqrt{1+4LA_k}}{2L} \\
A_{k+1} &= A_k + a_k \\
\tilde{x}_k &= \frac{A_k}{A_{k+1}} y_k + \frac{a_k}{A_{k+1}} x_k \\
y_{k+1} &= \operatorname*{argmin}_{x\in X} \left\{ l_f(x;\tilde{x}_k) + \frac{L}{2}\|x-\tilde{x}_k\|^2 \right\} \\
x_{k+1} &= x_k + a_k L(y_{k+1} - \tilde{x}_k)
\end{aligned}$$

(2) Set $k \leftarrow k+1$ and go to (1).

Proposition 2.9. *There exists a sequence of affine functions* $\{\gamma_k\}_{k\geq 0}$ *such that* $\gamma_k \leq f$ *and*

$$A_k f(y_k) \leq \min \left\{ A_k \Gamma_k(x) + \frac{1}{2}\|x-x_0\|^2 \right\} \tag{$(1)_k$}$$

$$x_k = \operatorname*{argmin}_{x\in\mathbb{R}^n} \left\{ A_k \Gamma_k(x) + \frac{1}{2}\|x-x_0\|^2 \right\} \tag{$(2)_k$}$$

where $\Gamma_k(x) = \left(\sum_{i=0}^{k-1} a_i \gamma_i(x)\right)/A_k$ *and* $\gamma_i = l_f(x;\tilde{x}_i)$.

[*****Aside**: It is important to know that if f is μ-strongly convex, then $\min f(x) \geq f_* + \frac{\mu}{2}\|x-x^*\|^2$. **This will show up on the exam!**]

Lemma 2.14. *For every* $k \geq 0$ *we have*

$$\begin{aligned}
A_k &= \sum_{i=0}^{k-1} a_i \\
A_{k+1}\Gamma_{k+1} &= A_k\Gamma_k + \alpha_k \gamma_k \\
\gamma_k &\leq f \\
A_k\Gamma_k &\leq A_k f
\end{aligned}$$

Proof. Trivial. □

Proof. [of previous proposition] We proceed by induction on k. The case for $k=0$ is obvious, so assume that it is true for k where $(1)_k$ and $(2)_k$ hold. In particular, using the previous lemma,

$$A_k\Gamma_k(x) + \frac{1}{2}\|x-x_0\|^2 \geq A_k f(y_k) + \frac{1}{2}\|x-x_k\|^2 \tag{3}$$

and so for all $x \in X$ (*) we have, using the lemma again, and letting $\tilde{x} = \tilde{x}(x) = \frac{A_k y_k + a_k x}{A_{k+1}}$,

$$\begin{aligned}
A_{k+1}\Gamma_{k+1}(x) + \frac{1}{2}\|x - x_0\|^2 &= A_k\Gamma_k(x) + a_k\gamma_k(x) + \frac{1}{2}\|x - x_0\|^2 \\
&\overset{(3)}{\geq} A_k f(y_k) + \frac{1}{2}\|x - x_k\|^2 + a_k\gamma_k(x) \\
&\geq A_k\gamma_k(x_k) + a_k\gamma_k(x) + \frac{1}{2}\|x - x_k\|^2 \\
&= A_{k+1}\gamma_k\left(\frac{A_k y_k + a_k x}{A_{k+1}}\right) + \frac{1}{2}\|x - x_k\|^2 \\
&= A_{k+1}\gamma_k(\tilde{x}) + \frac{1}{2}\left\|\frac{A_{k+1}}{a_k}(\tilde{x} - \tilde{x}_k)\right\|^2 \\
&= A_{k+1}\left(\gamma_k(\tilde{x}) + \frac{A_{k+1}}{2a_k^2}\|\tilde{x} - \tilde{x}_k\|^2\right) \\
&= A_{k+1}\left(\gamma_k(\tilde{x}) + \frac{L}{2}\|\tilde{x} - \tilde{x}_k\|^2\right) \\
&= A_{k+1}\left[l_f(\tilde{x}; \tilde{x}_k) + \frac{L}{2}\|\tilde{x} - \tilde{x}_k\|^2\right] \\
&\geq A_{k+1}\left[l_f(y_{k+1}; \tilde{x}_k) + \frac{L}{2}\|y_{k+1} - \tilde{x}_k\|^2\right]
\end{aligned}$$

since $\tilde{x}(x) - \tilde{x_k} = \frac{a_k}{A_{k+1}}(x - x_k)$. Hence $(1)_{k+1}$ follows. Next, for $(2)_{k+1}$, it is sufficient to show that

$$A_{k+1}\nabla\Gamma_{k+1} + x_{k+1} - x_0 = 0.$$

Directly, we have

$$\begin{aligned}
A_{k+1}\nabla\Gamma_{k+1} &= A_k\nabla\Gamma_k + a_k\nabla\gamma_k \\
&\overset{(2)_k}{=} x_0 - x_k + a_k\nabla\gamma_k \\
&= x_0 - x_{k+1}.
\end{aligned}$$

This is due to the construction of the algorithm:

$$\begin{aligned}
& y_{k+1} = \operatorname*{argmin}_{x \in X}\left\{\gamma_k(x) + \frac{L}{2}\|x - \tilde{x}_k\|^2\right\} \\
\implies & \nabla\gamma_k + L(y_{k+1} - \tilde{x}_k) = 0 \\
\implies & \nabla\gamma_k = L(\tilde{x}_k - y_{k+1}).
\end{aligned}$$

□

Remark 2.9. For the constrained case where we want (*) to become $x \in \mathbb{R}^n$, take

$$\gamma_k(x) = \langle L(\tilde{x}_k - y_{k+1}), x - y_{k+1}\rangle + l_f(y_{k+1}; \tilde{x}_k)$$

which has the property that

$$\begin{aligned}
\gamma_k(y_{k+1}) &= l_f(y_{k+1}; \tilde{x}_k) \\
\nabla\gamma_k &= L(\tilde{x}_k - y_{k+1}) \\
\min_{x \in \mathbb{R}^n}\left\{\gamma_k(x) + \frac{L}{2}\|x - \tilde{x}_k\|^2\right\} &= \min_{x \in X}\left\{l_f(x; \tilde{x}_k) + \frac{L}{2}\|x - \tilde{x}_k\|^2\right\}.
\end{aligned}$$

The proof can be constructed in the same manner as before.

Corollary 2.5. *For every $k \geq 0$ and $x^* \in X^*$ we have*

$$f(y_k) - f_* \leq \frac{1}{2A_k}\|x^* - x_0\|^2 = \frac{d_0^2}{2A_k}.$$

One can then show that $a_k \geq \frac{\lambda}{2} + \sqrt{\lambda A_k}$ for $\lambda = 1/L$ and hence

$$A_{k+1} \geq \left(\sqrt{A_k} + \frac{\sqrt{\lambda}}{2}\right)^2 \implies \sqrt{A_{k+1}} \geq \sqrt{A_k} + \frac{\sqrt{\lambda}}{2} \implies A_k \geq \frac{k^2\lambda}{4} = \frac{k^2}{4L}.$$

Proof. Since $\Gamma_k(x) \leq f(x)$, then

$$\begin{aligned} A_k f(y_k) &\leq A_k f(x) + \frac{1}{2}\|x - x_0\|^2, \forall x \in X \\ \implies A_k f(y_k) &\leq A_k f(x^*) + \frac{1}{2}\|x^* - x_0\|^2. \end{aligned}$$

Strongly Convex Case

Suppose we start with the following two assumptions

(A1) f is differentiable on X and for $L > 0$ we have $|\nabla f(x) - \nabla f(\tilde{x})| \leq L\|x - \tilde{x}\|$ for all $x, \tilde{x} \in X$

(A2) f is μ-strongly convex

We then have that (A1), (A2) imply that for $x, \tilde{x} \in X$,

$$l_f(\tilde{x}, x) + \frac{\mu}{2}\|x - \tilde{x}\|^2 \leq f(\tilde{x}) \leq l_f(\tilde{x}, x) + \frac{L}{2}\|x - \tilde{x}\|^2$$

□

Algorithm 3. *The **Nesterov Algorithm for μ-strongly convex functions** under (A1), (A2) is*

(0) Let $x_0 \in \mathbb{R}^n$ be given and set $y_0 = x_0, k = 0, A_0 = 0, \frac{1}{L} \leq \lambda \leq \frac{1}{L-\mu}$.

(1) Compute

$$\begin{aligned} \lambda_k &= (1 + \mu A_k)\lambda \\ a_k &= \frac{1 + \sqrt{\lambda_k^2 + 4\lambda_k A_k}}{2} \\ A_{k+1} &= A_k + a_k \\ \tilde{x}_k &= \frac{A_k}{A_{k+1}} y_k + \frac{a_k}{A_{k+1}} x_k \\ \hat{x}_k &= \mathcal{P}_X(\hat{x}_k) \\ y_{k+1} &= \operatorname*{argmin}_{x \in X} \left\{ l_f(x; \hat{x}_k) + \frac{1}{2\lambda}\|x - \hat{x}_k\|^2 + \frac{\mu}{2}\|x - \hat{x}_k\|^2 \right\} \\ x_{k+1} &= x_k - \frac{a_k}{1 + A_k\mu}\left[\frac{y_{k+1} - \tilde{x}_k}{\lambda} + \mu(y_{k+1} - x_k)\right] \end{aligned}$$

(2) Set $k \hookleftarrow k + 1$ and go to (1).

Note that $a_k^2 = (A_k + a_k)\lambda_k = A_{k+1}\lambda_k$.

Proposition 2.10. *Let $q(y)$ be a μ-strongly convex function such that $q \leq f$ on X. For $\lambda > 0$ and $\hat{x} \in \mathbb{R}^n$, define*

$$\hat{y} = \operatorname*{argmin}_{y \in X} \left\{ q(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

Then the function

$$\gamma(y) = q(\hat{y}) + \left\langle \frac{\hat{x} - \hat{y}}{\lambda}, y - \hat{y} \right\rangle + \frac{\mu}{2}\|y - \hat{y}\|^2$$

satisfies

(a) $\gamma(\hat{y}) = q(\hat{y})$

(b) $\hat{y} = \operatorname{argmin}_{y \in Y} \left\{ q(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}$.

(c) γ *is* μ*-strongly convex on* $\mathbb{R}^n$

(d) $\gamma \leq q$ *on* X *which implies* $\gamma \leq f$ *on* X

Aside (for the exam). If $\underline{\phi} \leq \min\{\phi(x)\}$ and ϕ is β-strongly convex, with $\bar{x} = \operatorname{argmin}_x \phi(x)$ then $\underline{\phi} + \frac{\beta}{2}\|x - \bar{x}\|^2 \leq \phi(x)$.

Aside (for the exam). If f is μ-strongly convex, then $\lambda f + \frac{1}{2}\|x - x_0\|^2$ is $(\lambda\mu + 1)$ strongly convex.

Proposition 2.11. *For every* $k \geq 0$ *define*

$$\begin{aligned} \Gamma_k(y) &= \frac{\sum_{i=0}^{k-1} a_i \gamma_i(y)}{A_k}, \forall y \in \mathbb{R}^n \qquad (1) \\ \implies A_k \Gamma_k &= A_{k-1}\Gamma_{k-1} + a_{k-1}\gamma_{k-1} \end{aligned}$$

where

$$\begin{aligned} \gamma_k(y) &= q_k(y_{k+1}) + \left\langle \frac{\hat{x}_k - y_{k+1}}{\lambda}, y - y_{k+1} \right\rangle + \frac{\mu}{2}\|y - y_{k+1}\|^2 \\ q_k(y) &= l_f(y; \hat{x}_k) + \frac{\mu}{2}\|y - \hat{x}\|^2. \end{aligned}$$

Then we have

(a) Γ_k *is* μ*-strongly convex*

(b) $\gamma_k \leq q_k \leq f$ *on* X

(c) $\Gamma_k \leq f$ *on* X

(d) $x_k = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ A_k \Gamma_k(x) + \frac{1}{2}\|x - x_0\|^2 \right\}$

(e) $A_k f(y_k) \leq \min\{A_k \Gamma_k(x) + \frac{1}{2}\|x - x_0\|^2\}$

Proof. (a) Obvious.

(b) Use the fact that $q_k \leq f$ on X and $\gamma_k \leq q_k$ on X follows from the previous proposition.

(c) $\Gamma_k \leq f$ on X follows from (1) and the fact that $\gamma_i \leq f$ on X

(d) and (e) By induction on k. For $k = 0$, it is obvious since $A_0 = 0$. First, assume that $(d)_k$ and (e_k) holds. Then for all $x \in \mathbb{R}^n$ we have

$$A_k \Gamma_k(x) + \frac{1}{2}\|x - x_0\|^2 \geq A_k f(y_k) + \frac{A_k \mu + 1}{2}\|x - x_k\|^2.$$

So,

$$
\begin{aligned}
&\min_{x\in\mathbb{R}^n}\left\{A_{k+1}\Gamma_{k+1}(x)+\frac{1}{2}\|x-x_0\|^2\right\}\\
=&\min_{x\in\mathbb{R}^n}\left\{A_k\Gamma_k(x)+a_k\gamma_k(x)+\frac{1}{2}\|x-x_0\|^2\right\}\\
\geq&\min_{x\in\mathbb{R}^n}\left\{A_kf_k(x)+\frac{A_k\mu+1}{2}\|x-x_k\|^2+a_k\gamma_k(x)\right\}\\
\geq&\min_{x\in\mathbb{R}^n}\left\{A_k\gamma_k(x)+\frac{A_k\mu+1}{2}\|x-x_k\|^2+a_k\gamma_k(x)\right\}\\
\geq&\min_{x\in\mathbb{R}^n}\left\{(A_k+a_k)\gamma_k\left(\underbrace{\frac{A_ky_k+a_kx}{A_k+a_k}}_{\tilde{x}}\right)+\frac{A_k\mu+1}{2}\|x-x_k\|^2\right\}\\
=&\min_{\tilde{x}\in\mathbb{R}^n}\left\{A_{k+1}\gamma_k(\tilde{x})+\frac{A_k\mu+1}{2}\cdot\frac{A_{k+1}^2}{a_k^2}\|\tilde{x}-\tilde{x}_k\|^2\right\}\\
=&A_{k+1}\min_{\tilde{x}\in\mathbb{R}^n}\left\{\gamma_k(\tilde{x})+\frac{\lambda_k}{2\lambda}\cdot\frac{A_{k+1}^2}{a_k^2}\|\tilde{x}-\tilde{x}_k\|^2\right\}\\
=&A_{k+1}\min_{\tilde{x}\in\mathbb{R}^n}\left\{\gamma_k(\tilde{x})+\frac{1}{2\lambda}\|\tilde{x}-\tilde{x}_k\|^2\right\}
\end{aligned}
$$

Now

$$
\begin{aligned}
f(y_{k+1})&\leq l_f(y_{k+1};\hat{x}_k)+\frac{L}{2}\|y_{k+1}-\hat{x}_k\|^2\\
&\leq l_f(y_{k+1};\hat{x}_k)+\frac{\mu}{2}\|y_{k+1}-\hat{x}_k\|^2+\frac{L-\mu}{2}\|y_{k+1}-\tilde{x}_k\|^2\\
&\leq q_k(y_{k+1})+\frac{1}{2\lambda}\|y_{k+1}-\tilde{x}_k\|^2\\
&=\min_{y\in X}\left\{q_k(y)+\frac{1}{2\lambda}\|y-\tilde{x}\|^2\right\}\\
&=\min_{y\in\mathbb{R}^n}\left\{\gamma_k(y)+\frac{1}{2\lambda}\|y-\tilde{x}\|^2\right\}
\end{aligned}
$$

and hence $\min_{x\in\mathbb{R}^n}\left\{A_{k+1}\Gamma_{k+1}(x)+\frac{1}{2}\|x-x_0\|^2\right\}\geq f(y_{k+1})$. Let us prove that

$$(d)_{k+1}\iff A_{k+1}\nabla\Gamma_{k+1}(x_{k+1})+x_{k+1}-x_0=0.$$

By $(d)_k$, $A_k\nabla\Gamma_k(x_k)+x_k-x_0=0$ and also

$$\nabla\Gamma_k(x)=\nabla\Gamma_k(\bar{x})+\mu(x-\bar{x}),\forall x,\bar{x}\in\mathbb{R}^n,\forall k\geq 1. \tag{i}$$

So,

$$
\begin{aligned}
&x_{k+1}-x_0+A_{k+1}\nabla\Gamma_{k+1}(x_{k+1})\\
=&x_{k+1}-x_0+A_{k+1}\left[\nabla\Gamma_{k+1}(x_k)+\mu(x_{k+1}-x_k)\right]\\
=&x_{k+1}-x_0+A_{k+1}\nabla\Gamma_{k+1}(x_k)+A_{k+1}\mu(x_{k+1}-x_k)\\
=&x_{k+1}-x_0+A_k\nabla\Gamma_k(x_k)+a_k\nabla\gamma_k(x_k)+\mu A_{k+1}(x_{k+1}-x_k)\\
\overset{(i)}{=}&(1+\mu A_{k+1})(x_{k+1}-x_k)+a_k\nabla\gamma_k(x_k)\\
=&-a_k\left[\frac{\tilde{x}_k-y_{k+1}}{\lambda}+\mu(x_k-y_{k+1})\right]+a_k\nabla\gamma_k(x_k)\\
=&0
\end{aligned}
$$

□

Corollary 2.6. *For all $k \geq 1$ and $x^* \in X^*$ we have*

$$f(y_k) - f_* \leq \frac{1}{2A_k}\|x_0 - x^*\|^2$$

Proof. (e) implies that

$$\begin{aligned} A_k f(y_k) &\leq A_k \Gamma_k(x^*) + \frac{1}{2}\|x^* - x_0\|^2 \\ &\leq A_k f(x^*) + \frac{1}{2}\|x^* - x_0\|^2 \end{aligned}$$

□

Proposition 2.12. *For every $k \geq 1$ we have*

$$A_k \geq \max\left\{\frac{k^2}{4L}, \frac{1}{L}\left(1 + \sqrt{\frac{\mu}{2L}}\right)^{2(k-1)}\right\}.$$

Proof. Note that we have

$$\begin{aligned} a_k &\geq \frac{\lambda_k}{2} + \sqrt{\lambda_k A_k} \\ A_{k+1} &= A_k + a_k \\ &= \frac{\lambda_k}{2} + \sqrt{\lambda_k A_k} + A_k \\ &= \left(\sqrt{A_k} + \sqrt{\frac{\lambda_k}{2}}\right)^2 + \frac{\lambda_k}{4} \\ &\geq \left(\sqrt{A_k} + \sqrt{\frac{A_k \mu \lambda}{2}}\right)^2 + \frac{\mu A_k \lambda}{4} \\ &= A_k \left[\left(1 + \sqrt{\frac{\mu\lambda}{2}}\right)^2 + \frac{\mu\lambda}{4}\right] \\ &\geq A_k \left(1 + \sqrt{\frac{\mu\lambda}{2}}\right)^2 \\ &\geq A_k \left(1 + \sqrt{\frac{\mu}{2L}}\right)^2 \end{aligned}$$

and hence

$$A_k \geq A_1 \left(1 + \sqrt{\frac{\mu}{2L}}\right)^{2(k-1)} = \lambda \left(1 + \sqrt{\frac{\mu}{2L}}\right)^{2(k-1)}.$$

The first part of the maximum is from the original Nesterov method. □

2.9 Quasi-Newton Methods

Quasi-Newton Method's General Scheme

(0) Let $x^0 \in \mathbb{R}^n$ and $H_0 \in \mathbb{R}^{n \times n}$ symmetric and $H_0 > 0$ be given.

(1) For $k = 0, 1, 2, ...$ set

$$\begin{aligned} d_k &= -H_k g_k \\ x_{k+1} &= x_k + \alpha_k d_k. \end{aligned}$$

Update H_k to obtain $H_{k+1} > 0$ and symmetric. Here, we want $H_k \sim [\nabla^2 f(x_k)]^{-1}$.

Motivation

Let

$$q_k = g_{k+1} - g_k$$
$$p_k = x_{k+1} - x_k.$$

Then,

$$q_k = \nabla^2 f(x_k) p_k + o(\|p_k\|)$$

and if f is quadratic then $q_k = \nabla^2 f(x_k) p_k$.

Secant Equation

$p_k = H_{k+1} q_k$ which comes from our above approximation.

Rank-One Updates (SR1)

$H_{k+1} = H_k + a_k z_k z_k^T$ where $a_k \in \mathbb{R}$ and $z_k \in \mathbb{R}^n$. We want

$$p_k = H_{k+1} q_k = H_k q_k + a_k (z_k^T q_k) z_k$$

and so z_k is proportional to $p_k - H_k q_k$. If we choose $z_k = p_k - H_k q_k$ then

$$1 = a_k (z_k^T q_k) = a_k (p_k - H_k q_k)^T q_k \implies a_k = \frac{1}{(p_k - H_k q_k)^T q_k}$$

and we are left with the update

$$H_{k+1} = H_k + \frac{(p_k - H_k q_k)^T (p_k - H_k q_k)^T}{(p_k - H_k q_k)^T q_k}$$

Rank-Two Updates

$H_{k+1} = H_k + +auu^T + bvv^T$ for $a \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$. The secant equation implies that

$$p_k = H_{k+1} q_k = H_k q_k + a(u^T q_k) u + b(v^T q_k) v.$$

If we choose $u = p_k$ and $v = H_k q_k$ and enforce that

$$a(p_k^T q_k) = 1 \implies a = \frac{1}{p_k^T q_k}$$
$$b(q_k^T H_k q_k) = -1 \implies b = -\frac{1}{q_k^T H_k q_k}$$

then we have the **Davidon-Fletcher-Powell** (DFP) method with the update

$$H_{k+1}^{DFP} = H_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{H_k q_k q_k^T H_k}{q_k^T H_k q_k}.$$

Lemma 2.15. *For $c, d \in \mathbb{R}^n$, we have $\|c\|\|d\| \geq |c^T d|$ and equality holds if and only if c, d are colinear.*

Theorem 2.7. *If $p_k^T q_k > 0$ for all $k \geq 0$ then all H_k's generated in the above way is positive definite and symmetric.*

Proof. We proceed by induction on k. For $k = 0$, it is obvious since $H_0 > 0$ assumption. Assume that $H_k > 0$ for some $k \geq 0$. Let $x \neq 0$ be given. Then,

$$x^T H_{k+1} x = x^T H_k x + \frac{(p_k^T x)^2}{(p_k^T q_k)} - \frac{(q_k^T H_k x)^2}{q_k^T H_k q_k}.$$

Let $c = H_k^{1/2}x$ and $d = H_k^{1/2}q_k$. Then

$$\begin{aligned} x^T H_{k+1} &= \|c\|^2 - \frac{(c^T d)^2}{\|d\|^2} + \frac{(p_k^T x)^2}{(p_k^T q_k)} \\ &= \frac{\|c\|^2\|d\|^2 - (c^T d)^2}{\|d\|^2} + \frac{(p_k^T x)^2}{(p_k^T q_k)} \geq 0 \end{aligned}$$

from the previous lemma.

Claim. $x^T H_{k+1} x > 0$

Proof. Assume for contradiction that $x^T H_{k+1} x = 0$. Then $p_k^T = 0$ and c, d are colinear. That is $x = \lambda q_k$ for $\lambda \neq 0$. Hence $0 = p_k^T x = \lambda q_k^T p_k \neq 0$ and $H_{k+1} > 0$ as required. □

Question 1. *How can we guarantee the following condition for $\alpha_k > 0$?*

$$\begin{aligned} 0 < q_k^T p_k &= (g_{k+1} - q_k)^T(\alpha_k d_k) \\ &= \alpha_k(g_{k+1}^T d_k - g_k^T d_k) \end{aligned}$$

***Solution*.** It is enough to enforce $g_{k+1}^T d_k > g_k^T d_k$. An example of such an inexact line search is the Wolfe-Powell line search with $0 < \sigma < \tau < 1$. In particular, it has the conditions

(1) $f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k \sigma g_k^T d_k$

(2) $g_{k+1} d_k \geq \tau g_k d_k > g_k d_k$

Sherman-Morrison Formula

Proposition 2.13. *Assume that $A = B + USV^T$ where $S \in \mathbb{R}^{m\times m}, A, B \in \mathbb{R}^{n\times n}$ non-singular and $U, V \in \mathbb{R}^{n\times m}$. If $P = S^{-1} + V^T S^{-1} U$ is non-singular then*

$$A^{-1} = B^{-1} - B^{-1}UP^{-1}V^T B^{-1}$$

Other Rank-Two Updates

We could try the following iteration scheme

$$x_{k+1} = x_k - \alpha_k B_k^{-1} g_k, B_k \approx \nabla^2 f(x_k)$$

where B_{k+1} is obtained from B_k by the following rank two formula ($B_k p_k = q_k$):

$$B_{k+1}^{BFGS} = B_k + \frac{q_k q_k^T}{q_k^T p_k} - \frac{B_k p_k p_k^T B_k}{p_k^T B_k p_k}.$$

We call this the **Broyden-Fletcher-Goldfarb-Shannon** (BFGS) update. Using inversion, we can use the **Sherman-Morrison** formula to get

$$H_{k+1}^{BFGS} = (B_{k+1}^{BFGS})^{-1} = H_k + \left(1 + \frac{q_k^T H_k q_k}{q_k^T p_k}\right)\frac{p_k p_k^T}{p_k^T q_k} - \frac{p_k p_k^T H_k + H_k q_k p_k^T}{q_k^T p_k}$$

where $A = B_{k+1}^{BFGS}, B = B_k$ and $U = [q_k, B_k p_k], V = U$ and

$$S = \begin{bmatrix} \frac{1}{p_k^T q_k} & 0 \\ 0 & -\frac{1}{p_k^T B_k p_k} \end{bmatrix}.$$

Broyden's Family of Algorithms

Let $\phi = \phi_k \in \mathbb{R}$. Then the method is defined as

$$\begin{aligned} H_{k+1}^{\phi} &= (1-\phi)H_{k+1}^{DFP} + \phi H_{k+1}^{BFGS} \\ &= \phi H_{k+1}^{DFP} + \phi v_k v_k^T \end{aligned}$$

where

$$v_k = (q_k^T H_k q_k)^{1/2} \left(\frac{p_k}{p_k^T q_k} - \frac{H_k q_k}{q_k^T H_k q_k} \right).$$

Theorem 2.8. *If $H_k > 0, p_k^T q_k > 0, \phi \geq 0$ then $H_{k+1}^{\phi} > 0$.*

Theorem 2.9. *If $f(x) = \frac{1}{2}x^T Q x - b^T x + c$ with $Q > 0$ then for every $k \geq 0$ such that $g_k \neq 0$ we have:*

(1) $H_{k+1}^{\phi} q_j = p_j$ for $j = 0, 1, ..., k$

(2) $p_j^T Q p_i = 0$ for $0 \leq i < j \leq k$

(3) $p_0, ..., p_k$ are nonzero

Hence, the method terminates in $m \leq n$ iterations. If $m = n$ then $H_n = Q^{-1}$.

Remark 2.10. Since $q_j = Qp_j$ then $H_{k+1} q_j = p_j \implies (H_{k+1}Q)p_j = q_j$ for $j = 0, 1, ..., k$ and so $H_{k+1}Q$ acts like an identity operator on a particular subspace. In particular, $(H_{k+1}Q)x = x$ for all $x \in [p_0, ..., p_k]$.

Theorem. *If $H_0 = I$ then the iterates generated by Broyden's Quasi-Newton method, with the exact line search method, are identical to those generated by the conjugate gradient method.*

Convergence Result for General f

Theorem 2.10. *Let $f : \mathbb{R}^n \to \mathbb{R} \in \mathcal{C}^2(\mathbb{R}^n)$ and $x_0 \in \mathbb{R}^n$ be such that*

(1) $S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ is bounded and convex

(2) $\nabla^2 f(x) > 0$ for all $x \in S$

Let $\{x_k\}$ be a sequence generated by the Broyden Quasi-Newton method

$$x_k = x_k - \alpha_k H_k^{\phi_k} g_k$$

where $\phi_k \in [0, 1]$ and $H_0 = I$ and α_k is chosen by the W-P rule and $\alpha_k = 1$ is the first attempted step size. Then,

$$\lim_{k\to\infty} x_k = x^*$$

superlinearly in the sense that

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$$

where x^* is the unique global minimum of f over S.

Limited Memory Quasi-Newton Methods

The general formula for a Quasi-Newton method is

$$\begin{aligned} \phi(H, p, q) &= H + \left(1 + \frac{q^T H q}{p^T q}\right) \frac{pp^T}{p^T q} - \left(\frac{pq^T H + Hqp^T}{p^T q} \right) \\ &= \left(I - \frac{pq^T}{p^T q}\right) H \left(1 - \frac{qp^T}{p^T q}\right) + \frac{pp^T}{p^T q} \end{aligned}$$

and in particular, $H_k^{BFGS} = \phi(H_{k-1}, p_{k-1}, q_{k-1})$. The idea for the limited memory variant is that we store the latest pairs (p_i, q_i) for $i = k-1, ..i, k-m$ and generate H_k recursively through the steps

1. $H = H_0^k$ (simple, say $H = I$)

2. For $i = k-m, ..., k-1$ set $H \hookleftarrow \phi(H, p_i, q_i)$

3. $H_k = H$

It turns out this scheme makes the calculation of $H_k g_k$ very easy and the intermediate H matrices simple as well. The following is a full description of the algorithm.

Algorithm 4. *(For computing $H_k g$)*

$u \hookleftarrow g_k$
for $i = k-1, ..., , k-m$
$\alpha_i \hookleftarrow \frac{p_i^T u}{p_i^T q_i}$
$u \hookleftarrow u - \alpha_i q_i$
end for
$r \hookleftarrow H_0^k u$
for $i = k-m, ..., k-1$
$\beta \hookleftarrow \frac{q_i^T r}{p_i^T q_i}$
$r \hookleftarrow r + (\alpha_i - \beta) p_i$
end for
$H_k g_k \hookleftarrow r$

3 Constrained Optimization

The standard constrained optimization problem in this section will be denoted by

$$\begin{aligned} (ECP)\ \min\ & f(x) \\ \text{s.t. } & h_i(x) = 0,\ i = 1, 2, ..., m, \\ & x \in \mathbb{R}^n, \\ & f, h_i \in \mathcal{C}^2(\mathbb{R}^n) \end{aligned}$$

Definition 3.1. We say that $x \in \mathbb{R}^n$ is a **regular point** of (ECP) if

$$\nabla h_1(x), ..., \nabla h_m(x)$$

are linearly independent (equivalently $\nabla h(x) = [\nabla h_1(x) ... \nabla h_m(x)]$ is full column rank).

Remark 3.1. If x is a regular point, the matrix

$$\nabla h(x)^T \nabla h(x) \in \mathbb{R}^{m \times m}$$

is nonsingular.

Theorem 3.1. *(Lagrange Multiplier Theorem - First order necessary optimality conditions) If x^* is a regular local minimum of (ECP), then there exists a unique (∃!) $\lambda^* \in \mathbb{R}^m$ such that*

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

More compactly, we have

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

Proof. (construction) There exists $\epsilon > 0$ such that

$$f(x) \geq f(x^*), \forall x \in \bar{B}(x^*; \epsilon) = S \text{ s.t. } h(x) = 0. \qquad (0)$$

Let $\alpha > 0$ be given and, for every $k \in \mathbb{N}$, let

$$x_k \in \operatorname*{argmin}_{x \in S} F_k(x) := f(x) + \frac{k}{2} \|h(x)\|^2 + \frac{\alpha}{2} \|x - x^*\|^2$$

where existence is guaranteed by the Weierstrass theorem.

Claim 3.1. $\lim_{k \to \infty} x_k = x^*$.

Proof. (of claim) For all k we have

$$F_k(x_k) \leq F_k(x^*) \iff f(x_k) + \frac{k}{2}\|h(x_k)\|^2 + \frac{\alpha}{2}\|x_k - x^*\|^2 \leq f(x^*). \qquad (1)$$

Since $f(x)$ is bounded on S, we have $\{f(x_k)\}$ is bounded. As $k \to \infty$, we have

$$\lim_{k\to\infty} \|h(x_k)\| = 0. \qquad (2)$$

Let $\bar{x}$ be an accumulation point of $\{x_k\}$. By (2), we have $h(\bar{x}) = 0$ and by (1) we have

$$f(\bar{x}) + \frac{\alpha}{2}\|\bar{x} - x^*\|^2 \leq f(x^*).$$

Since $\bar{x} \in S$ and $h(\bar{x}) = 0$, by (0), we have

$$f(\bar{x}) \geq f(x^*) \qquad (4)$$

and by (3),(4), $\|x - x\|^* = 0$. □

(Th. proof cont.) For all k sufficiently large, $x_k \in \text{int}(S)$ and hence $\nabla F_k(x_k) = 0$ and $\nabla^2 F_k(x_k) \geq 0$. Now,

$$\begin{aligned} 0 &= \nabla F_k(x_k) \\ &= \nabla f(x_k) + k\nabla h(x_k)h(x_k) + \alpha(x_k - x^*) \\ &= \nabla f(x_k) + \nabla h(x_k)\lambda_k + \alpha(x_k - x^*) \end{aligned}$$

where $\lambda_k = kh(x_k)$.
Claim 3.2. $\{\lambda_k\} \to \lambda^*$ for some $\lambda^* \in \mathbb{R}^m$.

Proof. We have

$$\begin{aligned} &\nabla h(x_k)^T \nabla h(x_k)\lambda_k = -\nabla h(x_k)^T \left[\nabla f(x_k) + \alpha(x_k - x^*)\right] \\ \implies &\lambda_k = -\left[\nabla h(x_k)^T \nabla h(x_k)\right]^{-1} \nabla h(x_k)^T \left[\nabla f(x_k) + \alpha(x_k - x^*)\right] \\ \implies &\lim_{k\to\infty} \lambda_k = -\left[\nabla h(x^*)^T \nabla h(x^*)\right]^{-1} \nabla h(x^*)^T \left[\nabla f(x^*)\right] := \lambda^* \end{aligned}$$

□

(Th. proof cont.) Taking limits with the above results gives

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

□

Theorem 3.2. *(Second Order Necessary Conditions) If x^* is a regular local minimum of (ECP), then there exists a unique $\lambda^* \in \mathbb{R}^m$ such that*

$$\nabla f(x^*) + \nabla h(x^*)\lambda^* = 0$$

and

$$d^T \left(\nabla^2 f(x^*) + \nabla^2 h(x^*)\lambda^*\right) d \geq 0$$

for all $d \in V(x^*)$ where

$$V(x^*) = \{d \in \mathbb{R}^n : \nabla h(x^*)^T d = 0\}.$$

Proof. Define

$$F_k(x) := f(x) + \frac{k}{2}\|h(x)\|^2 + \frac{\alpha}{2}\|x - x^*\|^2$$

and note for all k sufficiently large,

$$\begin{aligned} 0 &\leq \nabla^2 F_k(x_k) \\ &= \nabla^2 f(x_k) + \nabla^2 h_i(x_k)\lambda_k + k\nabla h(x_k)\nabla h(x_k)^T + \alpha I \end{aligned}$$

Let $d \in V(x^*)$ be given where $\nabla h(x^*)^T d = 0$ and define

$$\begin{aligned} d_k &= d - \nabla h(x_k) \left[\nabla h(x_k)^T \nabla h(x_k)\right]^{-1} \nabla h(x_k)^T d \\ &= \text{Proj}_{\text{Null}(\nabla h(x_k)^T)}(x_k). \end{aligned}$$

Note that $\nabla h(x_k)^T (x_k) = 0$ and $d_k \to d$ as $k \to \infty$. Hence, we get

$$0 \le d_k^T \left(\nabla f(x_k) + \nabla^2 h(x_k)\lambda_k\right) d_k + \alpha \|d_k\|^2$$

and as $k \to \infty$ we obtain

$$0 \le d^T \left(\nabla f(x^*) + \nabla^2 h(x^*)\lambda^*\right) d + \alpha \|d\|^2.$$

As $\alpha > 0$ is arbitrary, we take $\liminf_{\alpha > 0}$ on both sides and the result follows. □

Definition 3.2. The **Lagrangian function** $L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ is defined as

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

Remark 3.2. The necessary first order optimality condition is equivalent to $\nabla_x L(x^*, \lambda^*) = 0$ and feasibility is $\nabla_\lambda L(x^*, \lambda^*) = 0$. The necessary second order optimality condition is equivalent to $d^T \nabla^2_{xx} L(x^*, \lambda^*) d \ge 0$ for all $d \in V(x^*)$.

The sufficient second order condition is $d^T \nabla^2_{xx} L(x^*, \lambda^*) d > 0$ for all $0 \ne d \in V(x^*)$.

Theorem 3.3. *(Second Order Necessary Conditions) Assume that $f, h \in \mathcal{C}^2$ and x^* is a regular local minimum of (ECP). Then there exists $\lambda^* \in \mathbb{R}^m$ such that*

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

and

$$d^T \nabla^2_{xx} L(x^*, \lambda^*) d \ge 0$$

for all $d \in V(x^*) = \{d \in \mathbb{R}^n : \nabla h(x^*)^T d = 0\}$.

Theorem 3.4. *(Second Order Sufficient Conditions) Assume that $f, h \in \mathcal{C}^2$ and $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ is such that*

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) = 0, h(x^*) = 0, \\ d^T \nabla^2_{xx} L(x^*, \lambda^*) d > 0, \forall 0 \ne d \in V(x^*). \end{aligned}$$

Then x^* is a strictly local minimum of ECP. In fact, there exists $\gamma > 0, \epsilon > 0$ such that

$$f(x) \ge f(x^*) + \frac{\gamma}{2}\|x - x^*\|, \forall x \in \bar{B}(x^*, \epsilon) \text{ s.t. } h(x) = 0.$$

Proof. Define

$$L_c(x, \lambda) = f(x) + \lambda^T h(x) + \frac{c}{2}\|h(x)\|^2$$

for $c \in \mathbb{R}$. We have

$$\begin{aligned} \nabla_x L_c(x, \lambda) &= \nabla f(x) + \nabla h(x)[\lambda + ch(x)] \\ &= \nabla_x L(x, \lambda + ch(x)) \end{aligned}$$

and

$$\begin{aligned} \nabla^2_{xx} L_c(x, \lambda) &= \nabla^2 f(x) + \left[\sum_{i=1}^m (\lambda + ch(x))_i \nabla^2 h_i(x)\right] + c\nabla h(x) \nabla h(x)^T \\ &= \nabla^2_{xx} L(x, \lambda + ch(x)) + c\nabla h(x) \nabla h(x)^T. \end{aligned}$$

For $(x, \lambda) = (x^*, \lambda^*)$, we have

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

and

$$\nabla^2_{xx} L_c(x^*, \lambda^*) = \nabla^2_{xx} L(x^*, \lambda^*) + c\nabla h(x^*) \nabla h(x^*)^T.$$

Lemma 3.1. *Let P, Q be $n \times n$ symmetric matrices such that $Q \geq 0$ and $d^T P d > 0$ for every $d \neq 0$ such that $d^T Q d = 0$. Then $\exists \bar{c} \in \mathbb{R}$ such that*

$$P + cQ > 0, \forall c \geq \bar{c}.$$

Proof. Assume for contradiction that for all $k \in \mathbb{N}$, $\exists d_k \in \mathbb{R}^n$ such that $\|d_k\| = 1$ and

$$d_k^T (P + kQ) d_k \leq 0.$$

Without loss of generality, assume that $d_k \to d$. Then,

$$d^T P d + \limsup_{k \to \infty} k d_k^T Q d_k \leq 0 \implies d^T Q d = 0, d^T P d \leq 0, d \neq 0$$

which contradicts our assumptions. □

The application of the above lemma with $P = \nabla^2_{xx} L(x^*, \lambda^*)$ and $Q = \nabla h(x^*) \nabla h(x^*)^T$ implies that there is a sufficiently large $\bar{c} \in \mathbb{R}$ such that $\nabla^2_{xx} L_c(x^*, \lambda^*) > 0$ and $\nabla_x L_c(x^*, \lambda^*) = 0$ for any $c > \bar{c}$. So x^* is a strict local minimum of

$$\begin{aligned} &\min_x \; L_c(x, \lambda^*) \\ &\text{s.t. } x \in \mathbb{R}^n. \end{aligned}$$

In fact, there exists $\gamma > 0, \epsilon > 0$ such that

$$\begin{aligned} &L_c(x, \lambda^*) \geq L_c(x^*, \lambda^*) + \frac{\gamma}{2} \|x - x^*\|^2 \\ &\forall x \in \bar{B}(x^*; \epsilon). \end{aligned}$$

Since $L_c(x, \lambda) = f(x)$ for every x such that $h(x) = 0$, then if $x \in \bar{B}(x^*, \epsilon)$ and $h(x) = 0$ then

$$\begin{aligned} f(x) = L_c(x, \lambda^*) &\geq L_c(x^*, \lambda^*) + \frac{\gamma}{2} \|x - x^*\|^2 \\ &= f(x^*) + \frac{\gamma}{2} \|x - x^*\|^2. \end{aligned}$$

□

Theorem 3.5. *Let (x^*, λ^*) be a regular local minimum and Lagrange multiplier for (ECP) satisfying the 2nd order sufficiency condition. Then $\exists \delta > 0$ such that $\forall u \in \bar{B}(0, \delta)$ there exists a pair of regular local minimum and Lagrange multipliers $p(u) = (x(u), \lambda(u))$ for $(ECP)_u$ which is continuously differentiable,*

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

and

$$\nabla p(u) = -\lambda(u), p(u) = f(x(u)).$$

where $(ECP)_u$ is the problem

$$\begin{aligned} &\min f(x) \\ &\textit{s.t. } h(x) = u \end{aligned}$$

Note that $\nabla p(0) = -\lambda^*$.

3.1 General NLPs

Consider the problem

$$\begin{aligned} (NLP) \quad &\min f(x) \\ &\text{s.t. } h(x) = 0 \\ &\quad\;\; g(x) \leq 0 \end{aligned}$$

where $g = (g_1, ..., g_r) : \mathbb{R}^n \mapsto \mathbb{R}^r$.

Notation 2. For $x \in \mathbb{R}^n$, we let $A(x) = \{j : g_j(x) = 0\} \subseteq \{1, 2, ..., r\}$ and

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

Definition 3.3. We say $x \in \mathbb{R}^n$ is **regular** if

$$\begin{cases} \nabla h_i(x), & i = 1, ..., m \\ \nabla g_j(x), & j \in A(x) \end{cases}$$

are linearly independent.

Theorem 3.6. *(KKT [**Karush-Kuhn-Tucker**] Necessary Optimality Conditions)*
Let x^* be a regular local minimum of (NLP). Then $\exists!(\lambda^*, \mu^*) \in \mathbb{R}^m \times \mathbb{R}^r$ such that

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0,$$
$$\mu^* \geq 0, \mu_j = 0, \forall j \notin A(x^*).$$

If, in addition, $f, g, h \in \mathcal{C}^2$ then

$$d^T \nabla^2_{xx} L(x^*, \lambda^*) d \geq 0$$

for every $d \in V(x^*)$ where

$$V(x^*) = \left\{ d \in \mathbb{R}^n : {}_{\nabla g_j(x^*)^T d = 0, j \in A(x^*)}^{\nabla h(x^*)^T d = 0} \right\}.$$

Proof. Consider the (ECP)

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } h(x) = 0 \\ &\quad\; g_j(x) = 0, j \in A(x^*) \end{aligned}$$

where clearly x^* is a regular local minimum of (ECP) [**prove this as an exercise**]. By the necessary optimality conditions for (ECP), there exists unique $\lambda^* \in \mathbb{R}^m$ and $\{\mu_j^*\}_{j \in A(x^*)}$ such that

$$\nabla f(x^*) + \nabla h(x^*) \lambda^* + \sum_{j \in A(x^*)} \mu_j^* \nabla g_j(x^*) = 0.$$

The second order necessary conditions of (ECP) also translate directly to the second order conditions of (NLP), once we prove that $\mu \geq 0$. To do this, we define

$$F_k(x) = f(x) + \frac{k}{2} \|h(x)\|^2 + \frac{k}{2} \|g^+(x)\|^2 + \frac{\alpha}{2} \|x - x^*\|^2$$

where $\alpha > 0$ and $g_j^+(x) = \max(0, g_j(x))$. Let

$$\begin{aligned} x_k \in &\operatorname{argmin} F_k(x) \\ &\text{s.t. } x \in \bar{B}(x^*, \epsilon) \end{aligned}$$

where $\epsilon > 0$ is such that $f(x) \geq f(x^*)$ for all $x \in \bar{B}(x^*, \epsilon)$. Using similar arguments as before, $x_k \to x^*$. So,

$$\nabla F_k(x_k) = 0, \nabla^2 F_k(x_k) \geq 0$$

and hence

$$\nabla f(x_k) + \nabla h(x_k) \lambda^k + \nabla g(x_k) \mu^k + \alpha(x_k - x^*) = 0$$

where $\lambda^k = k \cdot h(x_k), \mu^k = k \cdot g^+(x_k)$. Now for k sufficiently large, $g_j(x_k) < 0$ for $j \notin A(x^*)$. and hence $g_j^+(x_k) = 0$ for $j \notin A(x^*)$ and so $\mu_j^k = 0$ for $j \notin A(x^*)$. It is easy to show

$$\lambda^k \to \lambda^*$$
$$\mu_j^k \to \mu_j^*, j \in A(x^*)$$

and as $\mu^k \geq 0$, $\mu^* \geq 0$ as well. □

Theorem 3.7. *(Second Order Sufficient Conditions) Assume* $f, g, h \in \mathcal{C}^2$ *and* $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r$ *satisfying*

$$\begin{aligned} &\nabla_x L(x^*, \lambda^*, \mu^*) = 0 \\ &h(x^*) = 0, g(x^*) \leq 0 \\ &\mu^* \geq 0 \\ &\mu_j^* = 0, j \notin A(x^*) \\ &d^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0 \end{aligned}$$

for all

$$\begin{aligned} d \neq 0 \\ \nabla h(x^*)^T d = 0 \\ g_j(x^*)^T d = 0, j \in A(x^*). \end{aligned}$$

Also assume that $\mu_j > 0$ *for* $j \in A(x^*)$. *Then* x^* *is a strict local minimum.*

Proof. Consider the (ECP)

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & h(x) = 0 \\ & g(x) + s^2 = 0. \end{aligned}$$

Clearly, x^* is a strict local minimum of (NLP) if and only if $(x^*, s^*) = (x^*, [-g(x^*)]^{1/2})$ is a strict local minimum of (ECP). The 1st order sufficiency conditions of (ECP) lead us to the existence of μ^*, λ^* such that

$$\begin{aligned} \nabla_x L(x^*, \lambda^*, \mu^*) = 0 \\ 2\mu_j^* s_j^* = 0, j = 1, 2, ..., r \\ h(x^*) = 0, g(x^*) + (s^*)^2 = 0 \end{aligned}$$

and the 2nd order conditions lead us to the existence of $(d, \hat{d}) \neq 0$ such that

$$\begin{aligned} \nabla h(x)^T d = 0 \\ \nabla g_j(x)^T d + 2 s_j \hat{d}_j = 0, j = 1, 2, ..., r \end{aligned} \implies d^T \nabla L^2_{xx}(x^*, \lambda^*, \mu^*) d + 2 \sum_{j=1}^r \mu_j^* (\hat{d}_j)^2 > 0.$$

Now,

$$2\mu_j^* s_j^* = 0 \iff 2\mu_j^* (-g_j(x^*))^{1/2} \iff \mu_j^* g_j(x^*) = 0$$

which follows from

$$\mu^* \geq 0, \mu_j^* = 0, j \notin A(x^*).$$

Next, let $(d, \hat{d}) \neq 0$ be given. Assume

$$\begin{aligned} &\nabla h(x)^T d = 0 \\ &\nabla g_j(x)^T d + 2 s_j \hat{d}_j = 0, j = 1, 2, ..., r. \end{aligned}$$

Then,

$$\begin{aligned} &\nabla h(x)^T d = 0 \\ &\nabla g_j(x)^T d = 0, j \in A(x^*). \end{aligned}$$

If $d \neq 0$ then we have

$$d^T \nabla L^2_{xx}(x^*, \lambda^*, \mu^*) d > 0$$

and hence

$$d^T \nabla L^2_{xx}(x^*, \lambda^*, \mu^*) d + 2 \underbrace{\sum_{j=1}^r \mu_j^* (\hat{d}_j)^2}_{\geq 0} > 0.$$

If $d = 0$ then we have $\hat{d} \neq 0$ and as long as

$$2 \sum_{j \in A(x^*)} \mu_j^* (\hat{d}_j)^2 > 0$$

then we are done. We generally assume that $\mu_j^* (\hat{d}_j)^2 \neq 0$ for some $j \in A(x^*)$. ☐

Proposition 3.1. *(Mangasarian-Fromovitz CQ) If $\nabla h_i(x^*) = 0$ and are linearly independent for $i = 1, 2, ..., m$ and $\exists d \in \mathbb{R}^m$ such that $\nabla h(x^*)^T d = 0, \nabla g_j(x^*)^T d < 0$ for $j \in A(x^*)$ then the first order necessary conditions are satisfied.*

Proof. (not proven in class) ☐

Proposition 3.2. *(Slater CQ) If h is affine, g_j is convex, and $\exists \bar{x}$ such that $g_j(\bar{x}) < 0$ for all $j \in A(x^*)$, then the previous proposition holds.*

Proof. Exercise. Use $d = \bar{x} - x^*$. ☐

Proposition 3.3. *(Linear/Concave CQ) If h is affine and g is concave, the first order necessary conditions hold without the regularity condition.*

Proof. (not proven in class) ☐

Proposition 3.4. *(General sufficiency condition) For the problem*

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & h(x) = 0 \\ & g(x) \leq 0 \\ & x \in X \end{aligned}$$

assume that (x^*, λ^*, μ^*) is such that x^* is feasible and

$$x^* \in \operatorname*{argmin}_{x \in X} L(x, \lambda^*, \mu^*)$$

with $\mu^* \geq 0$ and $(\mu^*)^T g(x^*) = 0$ where the second condition is equivalent to $\mu_j = 0$ for $j \notin A(x^*)$. Then x^* is a global minimum.

Note that if f, g are convex and h is affine, then $L(\cdot, \lambda^*, \mu^*)$ is convex and the previous statement is directly related to our previous sufficiency condition (convexity gives us a global minimum).

Proof. (not proven in class) ☐

3.2 Augmented Lagrangian Methods

Definition 3.4. For $c > 0$, the **augmented Lagrangian function** is defined as

$$L_c(x, \lambda) = f(x) + \lambda^T h(x) + \frac{c}{2} \|h(x)\|^2.$$

The **classical penalty approach** was

$$\min_{x \in X} f(x) + \frac{c_k}{2} \|h(x)\|^2 \text{ where } c_k \to \infty$$

and the modern approach is to use the augmented Lagrangian function.

Proposition 3.5. *Assume that $X = \mathbb{R}^n$ and (x^*, λ^*) is a pair satisfying the 2nd order sufficiency condition, i.e.,*

$$\begin{aligned} & \nabla_x L(x^*, \lambda^*) = 0, h(x^*) = 0 \\ & d^T \nabla^2_{xx} L(x^*, \lambda^*) d > 0 \text{ for every } d \text{ s.t. } \nabla h(x^*)^T d = 0. \end{aligned}$$

Then x^* is a strict local minimum of $L_c(\cdot, \lambda^*)$ for every c sufficiently large.

Example 3.1. Consider the problem

$$\min \frac{1}{2}(x_1^2 + x_2^2)$$
$$\text{s.t. } h(x) = x_1 - 1 = 0$$

where here $x^* = (1,0)$ and $\lambda^* = -1$. We also have (define)

$$L_c(x,\lambda) = \frac{1}{2}(x_1^2 + x_2^2) + \lambda(x_1 - 1) + \frac{c}{2}(x_1 - 1)^2$$
$$x(\lambda, c) = \operatorname*{argmin}_{x\in\mathbb{R}^n} L_c(x,\lambda) = \left(\frac{c-\lambda}{c+1}, 0\right)$$

for all $c > 0$. Now,

$$\lim_{\lambda\to\lambda^*} x(\lambda, c) = (1,0) = x^*.$$

Alternatively, for every $\lambda \in \mathbb{R}^n$,

$$\lim_{c\to\infty} x(\lambda, c) = (1,0) = x^*.$$

General Approach (Penalty)

For $\{c_k\} \subseteq \mathbb{R}_{++}$ and $\{\lambda_k\} \subseteq \mathbb{R}^n$, find $x_k \in \operatorname{argmin}_{x\in X} L_{c_k}(\cdot, \lambda_k)$.

Proposition 3.6. *(Quadratic Penalty Method) Assume that f, h are continuous, X is closed and (ECP) is feasible. Suppose $\{\lambda_k\}$ is bounded and $c_k \to \infty$. Then every limit point of $\{x_k\}$ is a global minimum of (ECP). Notationally, we may write $v^k = c_k$.*

Proof. Let $\bar{x}$ be a limit point of $\{x_k\}$. For all $x \in X$ and for all $k > 0$,

$$L_{c_k}(x_k, \lambda_k) \le L_{c_k}(x, \lambda_k) = f(x) + \lambda_k^T h(x) + \frac{c_k}{2}\|h(x)\|^2.$$

So if x is feasible for (ECP), then

$$L_{c_k}(x_k, \lambda_k) \le f(x), \forall k \ge 0$$

and hence for all $k \ge 0$,

$$L_{c_k}(x_k, \lambda_k) \le f_* := \inf_{h(x)=0, x\in X} f(x).$$

So

$$f(x_k) + \lambda_k^T h(x_k) + \frac{c_k}{2}\|h(x)\|^2 \le f_*, \forall k \ge 0.$$

Since $\{\lambda_k\}$ is bounded, there exists a subsequence $\{(x_k, \lambda_k)\} \overset{k\in K}{\to} (\bar{x}, \bar{\lambda})$. As $k \in K \to \infty$, we get

$$f(\bar{x}) + \bar{\lambda}^T h(\bar{x}) + \limsup_{k\in K} \frac{c_k}{2}\|h(x_k)\|^2 \le f_* \qquad (*)$$
$$\implies \|h(x_k)\| \overset{k\in K}{\to} 0$$
$$\implies h(\bar{x}) = 0$$

and since X is closed, $\bar{x} \in X$. So $(*)$ implies that $f(\bar{x}) \le f_*$ and hence $\bar{x}$ is a global minimum of (ECP). □

Proposition 3.7. *Assume that $X = \mathbb{R}^n$ and $f, g \in \mathcal{C}^1(\mathbb{R}^n)$. Assume also that*

$$\|\nabla_x L_{c_k}(x_k, \lambda_k)\| \le \epsilon_k$$

where $\{\lambda_k\}$ is bounded, $\epsilon_k \to 0$ and $c_k \to \infty$. Assume also $x_k \overset{k\in K}{\to} x^*$ where x^* is a regular point. Then there exists $\lambda^* \in \mathbb{R}^n$ such that

$$\lambda_k + c_k h(x_k) \to \lambda^*$$

and

$$\begin{cases} \nabla f(x^*) + \nabla h(x^*)\lambda^* = 0 \\ h(x^*) = 0. \end{cases}$$

Proof. Let $\bar{\lambda}_k = \lambda_k + c_k h(x_k)$. We have

$$\begin{aligned}\nabla_x L_{c_k}(x_k, \lambda_k) &= \nabla_x L(x_k, \lambda_k) + c_k \nabla h(x_k) h(x_k) \\ &= \nabla_x L(x_k, \bar{\lambda}_k) \\ &= \nabla f(x_k) + \nabla h(x_k) \bar{\lambda}_k\end{aligned}$$

which implies that

$$\bar{\lambda}_k = [\nabla h(x_k)^T \nabla h(x_k)]^{-1} \nabla h(x_k)^T [\nabla_x L_{c_k}(x_k, \lambda_k) - \nabla f(x_k)].$$

As $k \in K \to \infty$, we have

$$\bar{\lambda}_k \to -[\nabla h(x^*)^T \nabla h(x^*)]^{-1} \nabla h(x^*)^T \nabla f(x^*) =: \lambda^*$$

from regularity. Since $\bar{\lambda}_k \to \lambda^*$, we have $\{\bar{\lambda}_k\}$ is bounded. Since $\{\lambda_k\}$ is bounded, then $\{c_k h(x_k)\}$ is bounded and hence $h(x_k) \to 0$ since $c_k \to \infty$. By continuity, $h(x^*) = 0$. □

Hessian Ill-Conditioning

We have

$$Q_k = \nabla^2_{xx} L_{c_k}(x_k, \lambda_k) = \nabla^2_{xx} L(x_k, \bar{\lambda}_k) + c_k \nabla h(x_k) \nabla h(x_k)^T$$

and as $k \to \infty$,

$$\begin{aligned}&\nabla^2_{xx} L(x_k, \bar{\lambda}_k) \to \nabla^2_{xx} L(x^*, \lambda^*) \\ &\nabla h(x_k) \nabla h(x_k)^T \to \nabla h(x^*) \nabla h(x^*)^T\end{aligned}$$

and in the limit the matrix Q_k will have m eigenvalues tending to ∞ and $n - m$ eigenvalues which are bounded. So $\text{cond}(Q_k) \to \infty$.

Example 3.2. Consider the problem

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

where here $x^* = (1, 0)$ and $\lambda^* = -1$. We also have (define)

$$\begin{aligned}L_c(x, \lambda) &= \frac{1}{2}(x_1^2 + x_2^2) + \lambda(x_1 - 1) + \frac{c}{2}(x_1 - 1)^2 \\ \nabla_x L_c(x, \lambda) &= (x_1 + \lambda + c(x_1 - 1), x_2) \\ \nabla^2_{xx} L_c(x, \lambda) &= \begin{pmatrix} 1 + c & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\end{aligned}$$

Augmented Lagrangian Methods

Consider the augmented Lagrangian for (ECP), defined as

$$L_c(x, \lambda) = f(x) + \lambda^T h(x) + \frac{c}{2} \|h(x)\|^2$$

Recall that if (x^*, λ^*) is a pair satisfying the 2nd order sufficiency condition, then x^* is a strict local minimum of $L_c(\cdot, \lambda^*)$ for every $c \geq \bar{c}$.

Remark 3.3. Define $\{c_k\} \subseteq \mathbb{R}_{++}$ and $\{\lambda_k\} \subseteq \mathbb{R}^m$ and $x_k \in \operatorname{argmin}_{x \in X} L_{c_k}(x, \lambda_k)$. A previous proposition suggests the update $\lambda_{k+1} = \lambda_k + c_k h(x_k)$, which is called the **method of multipliers**.

Proposition 3.8. *Assume x^* is a regular local minimum of (ECP) which satisfies the 2nd order sufficiency condition. Let $\bar{c} \geq 0$ be such that*

$$\nabla^2 L_{\bar{c}}(x^*, \lambda^*) > 0.$$

Then $\exists \delta, \epsilon, M > 0$ such that *(a) For all (λ_k, c_k) satisfying*

$$\|\lambda_k - \lambda^*\| \le \delta c_k, c_k \ge \bar{c} \qquad (*)$$

the problem

$$\begin{aligned} &\min_x L_{c_k}(x, \lambda_k) \\ &\textit{s.t. } \|x - x^*\| < \epsilon \end{aligned}$$

has a unique global minimum x_k. Moreover,

$$\|x_k - x^*\| \le \frac{M}{c_k}\|\lambda_k - \lambda^*\|$$

(b) For all (λ_k, c_k) satisfying $(*)$,

$$\|\lambda_{k+1} - \lambda^*\| \le \frac{M}{c_k}\|\lambda_k - \lambda^*\|$$

where $\lambda_{k+1} = \lambda_k + c_k h(x_k)$.

Proof. (omitted) □

3.3 Global Method

A general algorithm is as follows:

(0) Let $\lambda_0 \in \mathbb{R}^m$ and $c_{-1} > 0$ be given and set $\epsilon_0 = \infty$ and $k = 0$.

(1) Set $c = c_{k-1}$.

(2) Compute $x \in \operatorname{argmin} L_c(\cdot, \lambda_k)$.
If $\|h(x)\| > \frac{1}{4}\epsilon_k$, set $c = 10c$ and go to (2).
Else, go to (3).

(3) Set $c_k = c, x_k = x, \lambda_{k+1} = \lambda_k + c_k h(x_k), \epsilon_{k+1} = \|h(x_k)\|$ and $k \hookleftarrow k+1$. Go to (1).

** Note that we may replace $\frac{1}{4}$ with any constant less than 1, and 10 with any constant greater than 1.

Proposition 3.9. *If the global method does not loop in (2), then every accumulation point x^* of $\{x_k\}$ which is regular satisfies*

$$\begin{aligned} &\nabla_x L(x^*, \lambda^*) = 0 \\ &h(x^*) = 0 \end{aligned}$$

for some $\lambda^* \in \mathbb{R}^m$. Moreover, λ^* is an accumulation point of $\{\lambda_k\}$.

Proof. We have

$$\|h(x_{k+1})\| \le \frac{1}{4}\|h(x_k)\| \implies h(x_k) \to 0 \implies h(x^*) = 0$$

and since λ_k is bounded and so is $c_k h(x_k)$ from the previous proposition, then $\lambda_{k+1} = \lambda_k + c_k h(x_k) \to \lambda^*$. □

Remark 3.4. If the method loops in (2), then the sequence of points $\{y^l\}$ generated satisfies

$$0 = \nabla_x L_{c_l}(y^l, \lambda_k) = \nabla f(y^l) + \nabla h(y^l)\left(\lambda_k + c_l h(y^l)\right).$$

If $y^l \overset{l \in L}{\to} y^*$ then $\nabla h(y^*) h(y^*) = 0, h(y^*) \ne 0$ and hence y^* is not regular. The fact that $\nabla_x L(x^*, \lambda^*) = 0$ follows from the fact that

$$\nabla_x L_{c_k}(x_k, \lambda_k) \to 0 \implies 0 = \nabla f(x^*) + \nabla h(x^*)\lambda^* = \nabla_x L(x^*, \lambda^*).$$

Remark 3.5. Consider the dual function $d_c(\lambda) = \min_{\|x - x^*\| \le \epsilon} L_c(x, \lambda)$. For 2nd order sufficient solutions, we have the following dual relationship:

$$\sup_{\lambda \in \mathbb{R}^m} d_c(\lambda) = f^* = \min \ f(x) \text{ s.t. } h(x) = 0, \|x - x^*\| \le \epsilon$$

Remark 3.6. The problem

$$(ICP)\ \min f(x) \quad \text{s.t. } g(x) \le 0$$

has equivalent (ECP) formulation

$$(E\tilde{C}P)\ \min f(x) \quad \text{s.t. } g(x) + u = 0, \quad u \in \mathbb{R}^m_+$$

for $(x,u) \in \mathbb{R}^n \times \mathbb{R}^m_+ = X$. Now define

$$\tilde{L}(x,u,\mu) = f(x) + \mu^T [g(x) + u] + \frac{c}{2}\|g(x) + u\|^2$$

and note that

$$\begin{array}{c}\min_{(x,u)} \tilde{L}(x,u,\mu) \\ \text{s.t. } (x,u) \in X\end{array} \equiv \begin{array}{c}\min_{x} L_c(x,\mu) \\ \text{s.t. } x \in \mathbb{R}^n\end{array}$$

where $L_c(x,\mu) = L_c(x, u(x,\mu), \mu)$ and

$$\begin{aligned} u(x,\mu) &= \text{argmin}\ \tilde{L}_c(x,u,\mu) \\ &= \operatorname*{argmin}_{u \ge 0} \mu^T u + \frac{c}{2}\|g(x) + u\|^2 \\ &= \max\left(-\frac{\mu}{c} - g(x), 0\right) \end{aligned}$$

Thus,

$$L_c(x,\mu) = f(x) + \mu^T g^+(x,\mu,c) + \frac{c}{2}\|g^+(x,\mu,c)\|$$

where $g^+(x,\mu,c) = \max(g(x), -\frac{\mu}{2})$. We update with $\mu_{k+1} = \max(0, \mu_k + c_k g(x_k))$ in the global method.

4 Barrier Methods

Consider the problem

$$(ICP)\ \min f(x) \quad \text{s.t. } g(x) \le 0, \quad x \in X$$

where $X \subseteq \mathbb{R}^n$ is closed, $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g : \mathbb{R}^n \mapsto \mathbb{R}^p$ is continuous. Let

$$\begin{aligned} \mathcal{F} &= \{x \in X : g(x) \le 0\} \\ \mathcal{F}^0 &= \{x \in X : g(x) < 0\} \end{aligned}$$

with the assumption that

(1) $\mathcal{F}^0 \ne \emptyset$

(2) $\mathcal{F} \subseteq \text{cl}\left(\mathcal{F}^0\right)$ (hence equality holds).

Barrier Function

This is a function $\psi : \mathbb{R}^p_{++} \mapsto \mathbb{R}$ continuous such that $\psi(y(x)) \to \infty$ as $x \to \text{bd}(\mathbb{R}^p_{++})$.

Barrier Subproblem

For $\mu > 0$, the subproblem is

$$\begin{aligned} &\min\ f(x) + \mu B(x) \\ &\text{s.t. } x \in \mathcal{F}^0 \end{aligned}$$

where $B(x) = \psi(-g(x))$.

Example 4.1.

(1) [Logarithmic]

$\psi(y) = -\sum_{i=1}^p \log y_i$ with $B(x) = -\sum_{i=1}^p \log(-g_i(x))$.

(2) [Inverse]

$\psi(y) = \sum_{i=1}^p \frac{1}{y_i}$ with $B(x) = -\sum_{i=1}^p \frac{1}{g_i(x)}$

Approach

For $\{\mu_k\} \subseteq \mathbb{R}_{++}$ such that $\mu_k \downarrow 0$, compute

$$x_k \in \operatorname*{argmin}_{x \in \mathcal{F}^0} f(x) + \mu_k B(x).$$

Theorem 4.1. *Every accumulation point of $\{x_k\}$ is an optimal solution of (ICP).*

Proof. Assume that $\bar{x} = \lim_{h \in K} x_k$ where clearly $\bar{x} \in \mathcal{F}$ since X is closed and g is continuous. There are two cases to consider.

(a) $\bar{x} \in \mathcal{F}^0$. In this case, $B(x_k) \to B(\bar{x})$ and also

$$f(x_k) + \mu_k B(x_k) \le f(x) + \mu_k B(x), \forall x \in \mathcal{F}^0. \qquad (*)$$

As $k \to \infty$ we have $f(\bar{x}) \le f(x), \forall x \in \mathcal{F}^0$ and since $\mathcal{F} \subseteq \text{cl}(\mathcal{F}^0)$ we have

$$f(\bar{x}) \le f(x), \forall x \in \mathcal{F}.$$

Hence $\bar{x}$ is an optimal solution.

(b) $\bar{x} \notin \mathcal{F}^0$. In this case, $B(x_k) \to \infty$ and there exists i such that $g_i(\bar{x}) = 0$. Hence, $B(x_k) \ge 0$ for all $k \in K$ sufficiently large and so by $(*)$,

$$f(x_k) \le f(x) + \mu_k B(x), \forall k \text{ sufficiently large.}$$

As $k \overset{k \in K}{\to} \infty$, we have use the same arguments in (a) to conclude that

$$f(\bar{x}) \le f(x), \forall x \in \mathcal{F}.$$

Hence $\bar{x}$ is an optimal solution. □

Logarithmic Barrier Method

Consider the problem (ICP) where $X = \mathbb{R}^n$. The **log barrier subproblem** is: for $\mu > 0$,

$$\begin{aligned} &\min_x\ f(x) - \mu \sum_{i=1}^p \log(-g_i(x)) = \phi_\mu(x) \\ &\text{s.t. } x \in \mathcal{F}^0. \end{aligned}$$

The optimality condition is

$$0 = \nabla \phi_\mu(x) = \nabla f(x) - \mu \sum_{i=1}^p \frac{\nabla g_i(x)}{g_i(x)}$$

or equivalently,

$$\begin{aligned} 0 &= \nabla f(x) + \sum_{i=1}^p \lambda_i \nabla g_i(x) \\ \lambda_i &= -\frac{\mu}{g_i(x)}, i = 1, ..., p. \end{aligned}$$

Recall that the necessary optimality conditions $(**)$ for (ICP) are

$$\begin{aligned} &\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0 \\ &\bar{\lambda}_i \geq 0, && i = 1, 2, ..., p \\ &\bar{\lambda}_i g_i(\bar{x}) = 0, && i = 1, ..., p. \end{aligned}$$

Theorem 4.2. *Assume that $\{x_k\}$ is a sequence of stationary points of $\min_{x \in \mathcal{F}^0} \phi_{\mu_k}(x)$ for some $\{\mu_k\} \downarrow 0$ and that $x_k \overset{k \in K}{\to} \bar{x}$ where $\bar{x}$ is a regular point of (ICP). Then*

$$\lambda_i^k = -\frac{\mu_k}{g_i(x_k)} \to \bar{\lambda}_i, i = 1, ..., p$$

for some $\bar{\lambda} \in \mathbb{R}^p$. Moreover, $(\bar{x}, \bar{\lambda})$ solves $(**)$.

Proof. For $k \in K$, we have

$$\begin{aligned} 0 &= \nabla f(x_k) - \mu_k \sum_{i=1}^{p} \frac{\nabla g_i(x_k)}{g_i(x_k)} \\ &= \nabla f(x_k) + \sum_{i=1}^{p} \lambda_i^k \nabla g_i(x_k) \end{aligned}$$

(1) $i \notin A(\bar{x})$. We have $g_i(\bar{x}) < 0 \implies \lambda_i^k = -\frac{\mu_k}{g_i(x_k)} \to 0$

(2) $i \in A(\bar{x})$. Then we have

$$\sum_{i \in A(\bar{x})} \lambda_i^k \nabla g_i(x_k) = -\nabla f(x_k) - \sum_{i \notin A(\bar{x})} \lambda_i^h \nabla g_i(x_k) \to -\nabla f(\bar{x}).$$

As before, using the fact that $\bar{x}$ is regular, we can show $\lambda_i^k \to \bar{\lambda}_i$. Hence,

$$\nabla f(x_k) - \sum_{i=1}^{p} \lambda_i^k \nabla g_i(x^k) \to \nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{\lambda}_i^k \nabla g_i(\bar{x}) = 0.$$

□

Lemma 4.1. *If u_k satisfies*

$$B^k u_k = b_k$$

and $B^k \to B$ which is full column rank. Then $u_k \to u$ for some u.

Proof. (**Exercise**) □

4.1 Interior Point Methods

Consider the standard LP problem

$$\begin{aligned} \min\ & c^T x = v^* \\ \text{s.t. } & Ax = b \\ & x \geq 0 \end{aligned}$$

with $X^0 = \{x > 0 : Ax = b\} \neq \emptyset$, A is $m \times n$, and $\text{rank}(A) = m$. Also assume that the set of optimal solutions X^* is non-empty. The log-barrier subproblem is: for $\mu > 0$

$$\begin{aligned} \min\ & c^T x - \mu \sum_{j=1}^{n} \log x_j \\ \text{s.t. } & Ax = b \\ & (x > 0). \end{aligned}$$

The optimality condition is

$$\begin{cases} c - \mu x^{-1} - A^T y = 0 \quad (x > 0) \\ Ax = b. \end{cases}$$

If we let $s = c - A^T y$ and $e = (1, 1, ..., 1)^T$ then the first condition is

$$s = \mu x^{-1} > 0 \implies x \circ s = \mu e$$

where $x \circ s$ is the Hadamard product. Now

$$b^T y \leq v^* \leq c^T x \implies c^T x - b^T y = x^T s = n\mu.$$

One can also show that

$$(y(\mu), s(\mu)) = (y, s) = \begin{aligned} \operatorname*{argmax}_{(\tilde{y}, \tilde{s})}\ & b^T \tilde{y} + \mu \sum_{i=1}^{n} \log \tilde{s}_i \\ \text{s.t. } & A^T \tilde{y} + \tilde{s} = c \\ & (\tilde{s} > 0) \end{aligned}$$

Proposition 4.1. *As $\mu \downarrow 0$ we have*

$$z(\mu) = (x(\mu), y(\mu), s(\mu)) \to (x^*, y^*, s^*).$$

The general algorithm is

(1) $z \approx z(\mu)$ approximation of $z(\mu)$

(2) Choose $\mu^+ < \mu$

(3) Obtain an approximation z^+ of $z(\mu^+)$

(4) Set $\mu \leftarrow \mu^+$ and go to step 1

Newton Step / Newton Direction

In the problem

$$\begin{aligned} \min\ & c^T x - \mu \sum_{j=1}^{n} \log x_j = \phi_\mu(x) \\ \text{s.t. } & Ax = b \\ & (x > 0). \end{aligned}$$

the **Newton step at** x is the subproblem

$$\begin{aligned} \min\ & \nabla \phi_\mu(x)^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 \phi_\mu(x) \Delta x \\ \text{s.t. } & A\Delta x = 0 \end{aligned}$$

which is equivalent to

$$\begin{aligned} \min\ & (c - \mu x^{-1})^T \Delta x + \frac{\mu}{2} \Delta x^T X^{-2} \Delta x \\ \text{s.t. } & A\Delta x = 0 \end{aligned}$$

where $X = \text{diag}(x)$ and $\Delta x = x^+ - x = \Delta x(x;\mu)$. The optimality conditions are

$$\begin{cases} c - \mu x^{-1} + \mu x^{-2}\Delta x - A^T y &= 0 \\ A\Delta x &= 0 \end{cases} \implies \begin{cases} x \circ s - \mu e + \mu x^{-1} \circ \Delta x &= 0 \\ A\Delta x &= 0 \end{cases}$$

where $y = y(x;\mu)$ is unique as the rows are A are linearly independent. If $\Delta x = 0$ then

$$\begin{cases} c - \mu x^{-1} - A^T y &= 0 \\ Ax &= b \\ s - \mu x^{-1} &= 0 \\ Ax &= b \\ A^T y + s &= c \end{cases} \implies \begin{cases} x &= x(\mu) \\ y &= y(\mu) \\ s &= s(\mu) \end{cases}.$$

<u>Closeness Criterion</u>

For $x \in X^0$ and $\mu > 0$, we define the closeness as

$$\delta_\mu(x) = \|x^{-1} \circ \Delta x(x;\mu)\| = \|x^{-1} \circ \Delta x\| = \frac{1}{\mu}\|x \circ s - \mu e\|$$

Proposition 4.2. *For $\mu > 0$ and $x \in X^0$ such that $\delta_\mu(x) < 1$, we have*

(a) $x^+ = x + \Delta x \in X^0$

(b) $s := s(x;t) > 0$ and (y,s) is strictly dual feasible

where $(\Delta x, y, s)$ are from the optimality conditions.

Proof. (a) Clearly

$$Ax^+ = A(x + \Delta x) = Ax + A\Delta x = b$$

so we have to show that $x^+ > 0$. We have

$$\begin{aligned} x^+ > 0 &\iff x + \Delta x > 0 \\ &\iff e + x^{-1}\Delta x > 0 \\ &\iff x^{-1}\Delta x > -e \\ &\iff \|X^{-1}\Delta x\|_\infty < 1 \\ &\iff \|X^{-1}\Delta x\| < 1 \\ &\iff \delta_\mu(x) < 1. \end{aligned}$$

(b) We have

$$1 > \delta_\mu(x) = \frac{1}{\mu}\|x \circ s - \mu e\| = \left\|\frac{xs}{\mu} - e\right\|$$

and as an **exercise**, one can show that this implies

$$\frac{xs}{\mu} > 0 \implies s > 0.$$

□

Proposition 4.3. *We have*

$$\|x \circ s - \mu e\| = \begin{aligned} \min_{(\tilde{y},\tilde{s})} \ & \|x \circ \tilde{s} - \mu e\| \\ \text{s.t. } & A^T\tilde{y} + \tilde{s} = c \end{aligned}$$

Proof. We may equivalently prove

$$\frac{1}{2}\|x \circ s - \mu e\|^2 = \begin{aligned} \min_{(\tilde{y},\tilde{s})} \ & \frac{1}{2}\|x \circ \tilde{s} - \mu e\|^2 \\ \text{s.t. } & A^T\tilde{y} + \tilde{s} = c \end{aligned}$$

which has optimality condition

$$\begin{aligned} x \circ (x \circ \hat{s} - \mu e) + \eta &= 0 \\ A\eta &= 0 \qquad (*) \\ A^T \hat{y} + \hat{s} &= c \end{aligned}$$

Since $(\hat{y}, \hat{s}, \eta) = (y, s, \mu\Delta x)$ satisfies $(*)$, the result follows. □

Proposition 4.4. *For $\mu > 0$ and $x \in X^0$ such that $\delta_\mu(x) < 1$ we have*

$$\delta_\mu(x^+) \leq \delta_\mu(x)^2$$

Proof. Let $s = s(x;\mu)$. Then,

$$\begin{aligned} x^+ \circ s - \mu e &= (x + \Delta x) \circ s - \mu e \\ &= x \circ s - \mu e + \Delta x \circ s \\ &= -\mu x^{-1} \circ \Delta x + s \circ \Delta x \\ &= (s - \mu x^{-1}) \circ \Delta x \\ &= (x \circ s - \mu e) \circ (x^{-1} \circ \Delta x) \\ &= -\mu (x^{-1} \circ \Delta x) \circ (x^{-1} \circ \Delta x) \end{aligned}$$

Hence,

$$\frac{1}{\mu}\|x^+ \circ s^+ - \mu e\| \leq \frac{1}{\mu}\|x^+ \circ s - \mu e\| \leq \|(x^{-1} \circ \Delta x) \circ (x^{-1} \circ \Delta x)\| \leq \|x^{-1} \circ \Delta x\|^2 = \delta_\mu(x)^2.$$

□

Remark 4.1. Define $\delta \in [\delta_\mu(x), 1)$ and the update step

$$\mu_+ = \left(1 + \frac{\gamma}{\sqrt{n}}\right)^{-1} \mu$$

and pick $\gamma > 0$ such that $(**)$ is satisfied below:

$$\delta_{\mu^+}(x) \overset{(*)}{\leq} \left[\left(1 + \frac{\gamma}{\sqrt{n}}\right)\delta_\mu(x) + \gamma\right] \leq \left[\left(1 + \frac{\gamma}{\sqrt{n}}\right)\delta_\mu(x) + \gamma\right] \overset{(**)}{\leq} \sqrt{\delta}.$$

where $(*)$ will be shown later. From the previous proposition,

$$\delta_\mu(x) \leq \delta \implies \delta_{\mu_+}(x) \leq \sqrt{\delta} \implies \delta_{\mu_+}(x^+) \leq \delta^2_{\mu_+}(x) \leq \delta$$

and so we have the invariant $\delta_\mu(x) \leq \delta$ with $x^+ = x + \Delta x(x;\mu_+)$. Let us prove $(*)$ above.

Proof. Let $s = s(x;\mu)$ and $y = y(x;\mu)$. Then,

$$\delta_\mu(x) = \frac{1}{\mu}\|x \circ s - \mu e\|.$$

Now

$$\begin{aligned}\delta_{\mu_+}(x) = & \min_{(\tilde{y},\tilde{s})} \frac{1}{\mu_+}\|x \circ \tilde{s} - \mu_+ e\| \\ & \text{s.t. } A^T\tilde{y} + \tilde{s} = c \\ & \le \frac{1}{\mu_+}\|x \circ s - \mu_+ e\| \\ & = \frac{1}{\mu_+}\|x \circ s - \mu e + (\mu - \mu_+)e\| \\ & \le \frac{1}{\mu_+}\|x \circ s - \mu e\| + (\mu - \mu_+)\|e\| \\ & \le \frac{1}{\mu^+}\left[\delta_\mu(x)\right] + (\mu - \mu_+)\sqrt{n}\end{aligned}$$

□

4.2 Interior Point Algorithm

(0) Let $(x_0, \mu_0) \in X^0 \times \mathbb{R}_{++}$ be such that $\delta_{\mu_0}(x_0) \le \delta$ and set $k \hookleftarrow 0$.

(1) Write $\mu_k > \frac{\epsilon}{n}\left(1 + \frac{\delta}{\sqrt{n}}\right)^{-1}$ and do:

$\mu_{k+1} = \mu_k \left(1 + \frac{\gamma}{\sqrt{n}}\right)^{-1}$ where γ is chosen to satisfy $(**)$

$x_{k+1} = x_k + \Delta x_k$ where $\Delta x_k = \Delta x(x_k, \mu_{k+1})$

Set $k \hookleftarrow k + 1$.

(2) Output x_k.

Proposition 4.5. *The algorithm terminates in $\mathcal{O}\left(\sqrt{n}\log\frac{n\mu_0}{\epsilon}\right)$ iterations with $x \in X^0$ such that $c^Tx - v^* \le \epsilon$.*

Proof. For every $k \ge 0$ we have $\delta_{\mu_k}(x_k) \le \delta, x_k \in X^0$. Let $(y_k, s_k) = (y(x_k,\mu_k), s(x_k,\mu_k))$. Then (y_k, s_k) is strictly dual feasible, so

$$\begin{aligned}c^Tx_k - v^* &\le c^Tx_k - b^Ty_k \\ &= x_k^T s_k \\ &= e^T(x_k \circ s_k) \\ &= e^T(x_k \circ s_k - \mu_k e + \mu_k e) \\ &= e^T(x_k \circ s_k - \mu_k e) + \mu_k n \\ &\le \|e\|\|x_k \circ s_k - \mu_k e\| + \mu_k n \\ &\le \sqrt{n}\delta_{\mu_k}(x_k) + \mu_k n \\ &\le \mu_k n\left(1 + \frac{\delta}{\sqrt{n}}\right).\end{aligned}$$

Assume that k is such that

$$\mu_k > \frac{\epsilon}{n\left(1 + \frac{\delta}{\sqrt{n}}\right)}$$

and note that $\mu_k = \mu_0 \left(1 + \frac{\gamma}{\sqrt{n}}\right)^{-k}$. So we have

$$\begin{aligned} \mu_0 \left(1 + \frac{\gamma}{\sqrt{n}}\right)^{-k} &> \frac{\epsilon}{n\left(1 + \frac{\delta}{\sqrt{n}}\right)} \\ \implies \frac{\mu_0 n \left(1 + \frac{\delta}{\sqrt{n}}\right)}{\epsilon} &> \left(1 + \frac{\gamma}{\sqrt{n}}\right)^{k} \\ \implies \log\left(\frac{\mu_0 n \left(1 + \frac{\delta}{\sqrt{n}}\right)}{\epsilon}\right) &> k \log\left(1 + \frac{\gamma}{\sqrt{n}}\right) \approx \frac{k\gamma}{\sqrt{n}} \\ \implies k &\leq \frac{\sqrt{n}}{\sqrt{\gamma}} \log\left(\frac{\mu_0 n \left(1 + \frac{\delta}{\sqrt{n}}\right)}{\epsilon}\right) \end{aligned}$$

using the fact that $\log(x) \geq \frac{x}{1+x}$. □

Remark 4.2. The optimality conditions can be re-written as

$$\begin{cases} Ax^2(c - \mu x^{-1}) - (Ax^2 A^T)y &= 0 \\ A\Delta x &= 0 \end{cases}$$

where this is a system of linear equations so that we can solve for $(y, \Delta x)$ to do the Newton step.

5 Duality

Consider the problem

$$\begin{aligned} (ICP) \ \min\ & f(x) \\ \text{s.t. } & g(x) \leq 0 \\ & x \in X \end{aligned}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g : \mathbb{R}^n \mapsto \mathbb{R}^r$. For $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^r$, we define the **Lagrangian function**

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

Definition 5.1. We say μ^* is a **geometric multiplier** for (ICP) if

$$\mu^* \geq 0 \text{ and } f_* = \inf_{x \in X} L(x, \mu^*).$$

Geometric Interpretation

Let $S = \{(g(x), f(x)) \in \mathbb{R}^{r+1} : x \in X\}$. We can see that (ICP) is equivalent to

$$\begin{aligned} \min\ & t \\ \text{s.t. } & (z, t) \in S \\ & z \leq 0 \end{aligned}$$

For $\mu \in \mathbb{R}^r$ and $c \in \mathbb{R}$, let $H(\mu, c) = \{(z, t) : z^T \mu + t = c\}$ be the hyperplane with normal $(\mu, 1)$ and its corresponding halfspace $H^+(\mu, c) = \{(z, t) : z^T \mu + t \geq c\}$.

Proposition 5.1. *We have*

$$S \subseteq H^+(\mu, c) \iff c \leq \inf_{x \in X} f(x) + \mu^T g(x) = \inf_{x \in X} L(x, \mu)$$

Proof. Directly,

$$\begin{aligned} & S \subseteq H^+(\mu, c) \\ \iff & g(x)^T\mu + f(x) \geq c, \forall x \in X \\ \iff & \inf_{x\in X} f(x) + \mu^T g(x) \geq c \end{aligned}$$

□

So for $\mu \in \mathbb{R}^r$,

$$f_* \geq \inf_{x\in X} f(x) + \mu^T g(x) = \max\left\{c : H^+(\mu, c) \supseteq S\right\}.$$

Proposition 5.2. *Let μ^* be a geometric multiplier. Then, x^* is a global minimum of (ICP) if and only if*

$$\begin{aligned} & x^* \in \operatorname*{argmin}_{x\in X} L(x, \mu^*) \\ & g(x^*) \leq 0 \\ & (\mu^*)^T g(x^*) = 0. \end{aligned}$$

Proof. ($\Longrightarrow$)Assume x^* is a global minimum of (ICP). Then $x^* \in X$, $g(x^*) \leq 0$ and $f_* = f(x^*)$. Hence

$$f_* \geq f(x^*) + (\mu^*)^T g(x^*) = L(x^*, \mu^*) \geq \inf_{x\in X} L(x, \mu^*) = f_*$$

where the last equality follows from the fact that μ^* is a geometric multiplier. So we must have

$$\begin{aligned} & (\mu^*)^T g(x^*) = 0 \\ & L(x^*, \mu^*) = \inf_{x\in X} L(x, \mu^*). \end{aligned}$$

($\Longleftarrow$) We have $x^* \in X, g(x^*) \leq 0$ and

$$f(x^*) = f(x^*) + (\mu^*)^T g(x^*) = L(x^*, \mu^*) = \inf_{x\in X} L(x, \mu^*) = f^*.$$

□

Remark 5.1. If f, g_j are convex for $j = 1, 2, ..., r$ and $X = \mathbb{R}^n$ then $L(\cdot, \mu^*)$ is convex and the above is reduced to: x^* is a global minimum of (ICP) if and only if $\nabla L(x^*, \mu^*) = 0$ if and only if

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

5.1 Dual Function

<u>ICP Duality</u>

Let us define $q : \mathbb{R}^r \mapsto [-\infty, \infty)$ as $q(\mu) = \inf_{x\in X} L(x, \mu)$. The **dual problem** is

$$\begin{aligned} q^* = & \sup_{\mu} q(\mu) \\ & \text{s.t. } \mu \geq 0. \end{aligned}$$

Proposition 5.3. *(ICP Weak Duality) For every $\mu \geq 0$ and $x \in X$ such that $g(x) \leq 0$ we have $f(x) \geq q(\mu)$ and hence $f^* \geq q^*$.*

Proof. Let $\mu \geq 0$ and $x \in X$ such that $g(x) \leq 0$ be given. Then,

$$f(x) \geq f(x) + \mu^T g(x) = L(x, \mu) \geq q(\mu).$$

□

Proposition 5.4. *Let $\mu^* \in \mathbb{R}^r$ be given. Then μ^* is a geometric multiplier if and only if $f^* = q^*$ and μ^* is a dual optimal solution.*

Proof. We note that μ^* is a geometric multiplier if and only if

$$f^* = q(\mu^*), \mu \geq 0 \iff f^* = q^* \text{ and } q^* = q(\mu^*)$$

from the fact that $f^* \geq q^* \geq q(\mu^*)$. □

Example 5.1. Consider the problem

$$\begin{aligned} \inf f(x) &= x \\ \text{s.t. } g(x) &= x^2 \leq 0 \\ x \in X &= \mathbb{R}. \end{aligned}$$

We have $x^* = 0, f^* = 0$. Now

$$q(\mu) = \inf_{x \in \mathbb{R}} x + \mu x^2 = \begin{cases} -\frac{1}{4\mu}, & \mu > 0 \\ -\infty, & \mu = 0 \end{cases} \implies \sup_{\mu \geq 0} q(\mu) = 0$$

but $\nexists \mu^*$ such that $q(\mu^*) = 0$ (i.e. the reverse direction of the previous proposition fails).

<u>NLP Duality</u>

For the (NLP) problem, define

$$\begin{aligned} L(x, \mu, \lambda) &= f(x) + \mu^T g(x) + \lambda^T h(x) \\ q(\mu, \lambda) &= \inf_{x \in X} L(x, \mu, \lambda) \end{aligned}$$

which are respectively the Lagrangian and dual function for (NLP).

Proposition 5.5. *(NLP Weak Duality) If x if feasible for (NLP) and $(\mu, \lambda) \in \mathbb{R}^r_+ \times \mathbb{R}^m$ then $f(x) \geq q(\mu, \lambda)$ and hence $f_* \geq q_*$,$f_* \geq q(\mu, \lambda)$, $f(x) \geq q_*$ where $q_* = \sup_{\mu \geq 0} q(\mu, \lambda)$.*

Proof. Let's compute $\inf_{x \in X} \sup_{(\mu,\lambda) \in \mathbb{R}^r_+ \times \mathbb{R}^m} L(x, \mu, \lambda)$. We have

$$\sup_{\substack{\mu \geq 0 \\ \lambda \in \mathbb{R}^m}} f(x) + \mu^T g(x) + \lambda^T h(x) = \begin{cases} f(x), & \text{if } g(x) \leq 0, h(x) = 0 \\ \infty, & \text{otherwise} \end{cases}.$$

So

$$\inf_{x \in X} \sup_{(\mu,\lambda) \in \mathbb{R}^r_+ \times \mathbb{R}^m} L(x, \mu, \lambda) = \sup_{\mu \geq 0} q(\mu, \lambda) \leq f(x).$$

□

Definition 5.2. The pair $(\mu^*, \lambda^*) \in \mathbb{R}^r \times \mathbb{R}^m$ is a **geometric multiplier** (G.M.) if $\mu^* \geq 0$ and $f_* = q(\mu^*) = q_*$.

Proposition 5.6. *Let $(\mu^*, \lambda^*) \in \mathbb{R}^r \times \mathbb{R}^m$ be given such that $\mu^* \geq 0$. Then, (μ^*, λ^*) is a G.M. if and only if (μ^*, λ^*) is a dual optimal solution and $f_* = q_*$.*

Proposition 5.7. *A pair $(x^*, (\mu^*, \lambda^*))$ is an optimal solution-G.M. pair if and only if*

$$\begin{aligned} &x \text{ is feasible} \\ &x^* \in \operatorname*{argmin}_{x \in X} L(x, \mu^*, \lambda^*) \\ &\mu^* \geq 0 \\ &g(x^*) \leq 0 \\ &(\mu^*)^T g(x^*) = 0. \end{aligned}$$

Proof. Similar to the ICP proof. □

Fact 5.1. *For $x \in X$ and $\mu \geq 0$ we have*

$$q(\mu, \lambda) \leq L(x, \mu, \lambda) \leq f(x).$$

Fact 5.2. *For $x \in X$ and $\mu \geq 0$ we have*

$$\sup_{\substack{\mu \geq 0 \\ \lambda \in \mathbb{R}^m}} L(x, \mu, \lambda) = \begin{cases} f(x), & \text{if } g(x) \leq 0, h(x) = 0 \\ \infty, & \text{otherwise} \end{cases}.$$

Proposition 5.8. *(Saddle Point) A pair $(x^*, (\mu^*, \lambda^*))$ is an optimal solution-G.M. pair if and only if*

$$x^* \in X, \mu \geq 0$$
$$L(x,^* \mu, \lambda) \leq L(x^*, \mu^*, \lambda^*) \leq L(x, \mu^*, \lambda^*), \forall (\mu, \lambda) \in \mathbb{R}^r_+ \times \mathbb{R}^m,$$
$$\forall x \in X$$

Proof. A pair $(x^*, (\mu^*, \lambda^*))$ is an optimal solution-G.M. pair if and only if

$$\begin{aligned} &x^* \in X, g(x^*) \leq 0, h(x^*) = 0 \\ &\mu^* \geq 0 \\ &f(x^*) = q(\mu^*, \lambda^*) \end{aligned}$$

if and only if

$$\begin{aligned} &x^* \in X, g(x^*) \leq 0, h(x^*) = 0 \\ &\mu^* \geq 0 \\ &f(x^*) = q(\mu^*, \lambda^*) = q(\mu^*, \lambda^*) \end{aligned}$$

if and only if

$$\begin{aligned} &x^* \in X, \mu^* \geq 0 \\ &\sup_{\substack{\mu \geq 0 \\ \lambda \in \mathbb{R}^m}} L(x^*, \mu^*, \lambda^*) = L(x^*, \mu^*, \lambda^*) = \inf_{x \in X} L(x^*, \mu^*, \lambda^*). \end{aligned}$$

□

5.2 Existence of G.M.'s

Here, let us consider the (NLP) problem

$$\begin{aligned} f_* = \inf\ & f(x) \\ \text{s.t. } & h(x) = 0 \\ & g(x) \leq 0 \\ & x \in X. \end{aligned}$$

Definition 5.3. X is **polyhedral** if $\exists D \in \mathbb{R}^{p \times n}$, $d \in \mathbb{R}^p$ such that $X = \{x \in \mathbb{R}^n : Dx \leq d\}$.

Proposition 5.9. *Assume that:*

* $f_* \in \mathbb{R}$

* *h, g are affine*

* *$f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex*

* *X is polyhedral*

Then (NLP) has a G.M. and as a consequence $f_* = q_*$.

Proposition 5.10. *Assume that:*

* $f_* \in \mathbb{R}$

* *h, g are affine*

* *$f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex quadratic*

* *X is polyhedral*

Then (NLP) has an optimal solution-G.M. pair.

General Case

Consider the general problem

$$\begin{aligned} f_* = \inf\ & f(x) \\ \text{s.t. } & Ax \le b \\ & g(x) \le 0 \\ & x \in X \end{aligned}$$

Proposition 5.11. *Assume that:*

* $f_* \in \mathbb{R}$

* *$X = C \cap P$ where P is polyhedral, C is convex*

* *$f : \mathbb{R}^n \mapsto \mathbb{R}, g_j : C \mapsto \mathbb{R}$ are convex*

* *$\exists \bar{x}$ such that $g(\bar{x}) < 0$, $A\bar{x} \le b$, and $\bar{x} \in ri(C) \cap P$*

Then (NLP) has a G.M. pair and as a consequence $f_* = q_*$.

Example 5.2. The problem

$$\begin{aligned} f_* = \min\ & e^{-\sqrt{x_1 x_2}} \\ \text{s.t. } & x_1 \le 0 \\ & (x_1, x_2) \ge 0 \end{aligned}$$

has $f_* = 1$ but for $\mu \ge 0$ we have

$$q(\mu) = \inf_{\substack{x_1 \ge 0 \\ x_2 \ge 0}} e^{-\sqrt{x_1 x_2}} + \mu x_1 = 0.$$

Duality Continued

Consider the primal-dual problem pair

$$\begin{aligned} \min\ & c^T x \\ \text{s.t. } & Ax \ge b \end{aligned}, \quad \begin{aligned} \max\ & b^T y \\ \text{s.t. } & A^T y \\ & y \ge 0. \end{aligned}$$

The dual function approach is equivalent to the dual problem above:

$$\begin{aligned} \max\ & d(\mu) \\ \text{s.t. } & \mu \ge 0 \end{aligned} = \begin{aligned} \max\ & b^T \mu \\ \text{s.t. } & A^T \mu = b \\ & \mu \ge 0 \end{aligned}$$

where

$$\begin{aligned} d(\mu) &= \inf_{x \in \mathbb{R}^n} c^T x + \mu^T(-Ax + b) = L(x, \mu) \\ &= \inf_{x \in \mathbb{R}^n} \left(c - A^T \mu\right)^T x + \mu^T b \\ &= \begin{cases} \mu^T b, & \text{if } c - A^T \mu = 0 \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Now consider the problem

$$\begin{aligned} \min\ & c^T x \\ \text{s.t. } & b - Ax = 0 \\ & x \geq 0 \end{aligned}$$

The dual function approach is equivalent to:

$$\begin{aligned} \max\ & d(\lambda) \\ \text{s.t. } & \lambda \in \mathbb{R}^m \end{aligned} = \begin{aligned} \max\ & b^T \lambda \\ \text{s.t. } & A^T \lambda \leq c \end{aligned}$$

where

$$\begin{aligned} d(\mu) &= \inf_{x \geq 0} c^T x + \lambda^T (b - Ax) = L(x, \mu) \\ &= \inf_{x \in \mathbb{R}^n} \left(c - A^T \lambda\right)^T + \lambda^T b \\ &= \begin{cases} \lambda^T b, & \text{if } c - A^T \lambda \geq 0 \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Both cases give us an intuition on how dual problems are constructed (in the linear case).

In the quadratic case, consider the problem

$$\begin{aligned} \min\ & c^T x + \frac{1}{2} x^T Q x \\ \text{s.t. } & Ax \geq 0 \end{aligned}$$

The dual function approach is equivalent to:

$$\begin{aligned} \max\ & d(\mu) \\ \text{s.t. } & \mu \geq 0 \end{aligned} = \begin{aligned} \max\ & \left(c - A^T \mu\right)^T x + \mu^T b + \frac{1}{2} x^T Q x \\ \text{s.t. } & c - A^T \mu + Qx = 0 \\ & \mu \geq 0 \end{aligned}$$

$$= \begin{aligned} \max\ & \mu^T b - \frac{1}{2} x^T Q x \\ \text{s.t. } & c - A^T \mu + Qx = 0 \\ & \mu \geq 0 \end{aligned}$$

where

$$\begin{aligned} d(\mu) &= \inf_{x \in \mathbb{R}^n} c^T x + \mu^T (-Ax + b) + \frac{1}{2} x^T Q x = L(x, \mu) \\ &= \inf_{x \in \mathbb{R}^n} \left(c - A^T \mu\right)^T x + \mu^T b + \frac{1}{2} x^T Q x \\ &= \begin{cases} \mu^T b - \frac{1}{2} x^T Q x, & \text{if } c - A^T \mu + Qx = 0 \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

and the condition arises from solving $\nabla d(\mu) = 0$. If Q is invertible, we have $x = Q^{-1}(A^T \mu - c)$ and so problem becomes

$$\begin{aligned} \max\ & \mu^T b - \frac{1}{2}(A^T \mu - c) Q^{-1} (A^T \mu - c) \\ \text{s.t. } & \mu \geq 0. \end{aligned}$$

5.3 Augmented Lagrangian Method vs. Duality

Consider the problem

$$\begin{aligned} f_* = \inf\ & f(x) && f : \mathbb{R}^n \mapsto \mathbb{R} \\ \text{s.t. } & Ax = b, && A \text{ is } m \times n \\ & x \in X && X \subseteq \mathbb{R}^n \end{aligned}$$

the **value function** is

$$\begin{aligned} v(u) = \inf_x\ & f(x) \\ \text{s.t. } & Ax - b = u \end{aligned}$$

where clearly, $v(0) = f_*$.

Proposition 5.12. *If X is convex and f is convex on X then $v(\cdot)$ is convex.*

Proof. Let $\lambda \in (0, 1)$ and $u_1, u_2 \in \mathbb{R}^n$ such that $v(u_i) < \infty$ for $i = 1, 2$ be given. We have

$$\begin{aligned} v(\lambda u_1 + (1-\lambda)u_2) &= \begin{array}{l} \inf\ f(x) \\ \text{s.t. } Ax - b = \lambda u_1 + (1-\lambda)u_2 \\ \quad x \in X \end{array} \\ &\leq \begin{array}{l} \inf\ f(x) \\ \text{s.t. } x = \lambda x_1 + (1-\lambda)x_2 \\ \quad Ax_1 - b = u_1, x_1 \in X \\ \quad Ax_2 - b = u_2, x_1 \in X \end{array} \\ &= \begin{array}{l} \inf\ f(\lambda x_1 + (1-\lambda)x_2) \\ \text{s.t. } Ax_1 - b = u_1, x_1 \in X \\ \quad Ax_2 - b = u_2, x_1 \in X \end{array} \\ &\leq \begin{array}{l} \inf\ \lambda f(x_1) + (1-\lambda) f(x_2) \\ \text{s.t. } Ax_1 - b = u_1 \\ \quad Ax_2 - b = u_2 \end{array} \\ &= \lambda v(u_1) + (1-\lambda) v(u_2). \end{aligned}$$

□

The dual problem to our original problem is

$$\begin{aligned} d(\lambda) &= \inf_{x \in X} f(x) + \lambda^T(b - Ax) = L(x, \lambda) \\ &= \inf_{u \in \mathbb{R}^m} \begin{pmatrix} \inf\ f(x) + \lambda^T(b - Ax) \\ \text{s.t. } Ax - b = u \\ \quad x \in X \end{pmatrix} \\ &\inf_{u \in \mathbb{R}^m} \left(v(u) - \lambda^T u \right) \end{aligned}$$

and so

$$-d(\lambda) = \sup_{u \in \mathbb{R}^m} \lambda^T u - v(u) =: v^*(\lambda)$$

where we call v^* the conjugate function of v. Note that $d(\lambda)$ is concave but usually not smooth.

Now note that the original problem is equivalent to

$$\begin{aligned} f_* = \inf\ & f(x) + \frac{\rho}{2}\|Ax - b\|^2 = f_\rho(x) \\ \text{s.t. } & Ax = b \\ & x \in X \end{aligned}$$

which has the dual function

$$v_\rho(u) = \begin{array}{c} \inf\ f_\rho(x) \\ \text{s.t. } Ax - b = u \\ x \in X \end{array}$$

with $v_\rho(0) = f_*$ and $v_\rho(u) \geq v(u)$.

Proposition 5.13. *If X is convex and f is convex on X then $v_\rho(\cdot)$ is ρ-strongly convex.*

Proof. We have

$$\begin{aligned} v_\rho(u) &= \begin{array}{c} \inf\ f_\rho(x) \\ \text{s.t. } Ax - b = u \\ x \in X \end{array} \\ &= \begin{array}{c} \inf\ f(x) + \frac{\rho}{2}\|u\|^2 \\ \text{s.t. } Ax - b = u \\ x \in X \end{array} \\ &= v(u) + \frac{\rho}{2}\|u\|^2 \end{aligned}$$

and since v is convex the result holds. The new dual problem, using the same steps as above, is

$$d_\rho(\lambda) = L_\rho(x, \lambda) = \inf_{u \in \mathbb{R}^m} v_\rho(u) - \lambda^T u = \inf_{u \in \mathbb{R}^m} v(u) - \lambda^T u + \frac{\rho}{2}\|u\|^2.$$

□

Proposition 5.14. *Assume that X is convex compact and f is convex on X. Then:*

(1) $d_\rho(\cdot)$ is concave and differentiable everywhere

(2) $\nabla d_\rho(\cdot)$ is $\frac{1}{\rho}$-Lipschitz continuous

(3) $\nabla d_\rho(\lambda) = -u_\rho(\lambda)$ where $u_\rho(\lambda) = \operatorname{argmin}_{u \in \mathbb{R}^m} v_\rho(u) + \lambda^T u$.

Recall the augmented Lagrangian method:

(0) $\lambda_0 \in \mathbb{R}^m$ is given; set $k \hookleftarrow 1$.

(1) Set $x_k = \operatorname{argmin}_{x \in X} L_\rho(x, \lambda_{k-1})$

(2) Set $\lambda_k = \lambda_{k-1} + \rho(b - Ax_k)$

(3) Set $k \hookleftarrow k + 1$ and go to (1).

Note that in step (2) we have

$$\lambda_k = \lambda_{k-1} + \rho \nabla d(\lambda_{k-1}) = \lambda_{k-1} + \frac{1}{L_\rho} \nabla d(\lambda_{k-1})$$

so this is steepest ascent on $d(\lambda_{k-1})$. Note that this step can be then replaced with

$$\lambda_k = \lambda_{k-1} + \frac{\theta}{L_\rho} \nabla d(\lambda_{k-1}) = \lambda_{k-1} + \theta\rho(b - Ax_k), \theta \in (0, 2)$$

Appendix

Definition 5.4. A coercive function f is a function where $\|x_n\| \to \infty$ implies that $f(x_n) \to \infty$.

Proposition 5.15. *A function is coercive if and only if for any $\alpha \in \mathbb{R}$, the set $\{x : f(x) \leq \alpha\}$ is compact.*

Proposition 5.16. *A coercive function has at least one global minimum, and the global minimum will be among the critical points of the function.*