Optimization Algorithm Notes

Depu Meng

June 3, 2019

Contents

1 Introduction to Optimization Algorithms

1.1 Goal of the Course

- Understand foundations of optimization
- Learn to analyze widely used optimization algorithms
- Be familiar with implementation of optimization algorithms

1.2 Basic Concepts

1.2.1 Problem Definition

Find the value of the decision variable s.t. objective function is maximized/minimized under certain conditions.

$$\begin{aligned} &\min f(x) \\ &s.t. x \in \mathcal{S} \subset \mathbb{R}^n \end{aligned} \tag{1.1}$$

Here, we call $\mathcal{S}$ *feasible region*.

We often denote constrained optimization Problem as

$$\begin{aligned} &\min f(x) \\ s.t. \quad &g_i(x) \geq 0, i = 1, ..., n \\ &b_i(x) = 0, i \in 1, ..., m \end{aligned} \tag{1.2}$$

Definition 1.2.1. *Global Optimality. For global optimal value $x^* \in \mathcal{S}$,*

$$f(x^*) \leq f(x), \forall x \in \mathcal{S} \tag{1.3}$$

Definition 1.2.2. *Local Optimality. For local optimal value $x^* \in \mathcal{S}$, $\exists U(x^*)$, such that*

$$f(x^*) \leq f(x), \forall x \in \mathcal{S} \cap U(x^*) \tag{1.4}$$

Definition 1.2.3. *Feasible direction. Let $x \in \mathcal{S}$, $d \in \mathbb{R}^n$ is a non-zero vector. if $\exists \delta > 0$, such that*

$$x + \lambda d \in \mathcal{S}, \forall \lambda \in (0, \delta) \tag{1.5}$$

*Then d is a **feasible direction** at x. We denote $F(x, \mathcal{S})$ as the set of feasible directions at x.*

Definition 1.2.4. *Descent direction.* $f(x) : \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$, d *is a non-zero vector. If* $\exists \delta > 0$, *such that*

$$f(x + \lambda d) < f(x), \forall \lambda \in (0, \delta) \tag{1.6}$$

Then d *is a* ***descent direction*** *at* x. *We denote* $D(x, f) = \{d | \bigtriangledown f(x)^T d < 0\}$ *as the set of descent direction at* x.

1.3 Optimal Conditions

1.3.1 Unconstrained Optimization

First-order necessary condition: $f(x)$ is differentiable at x,

$$\bigtriangledown f(x) = 0 \tag{1.7}$$

Second-order necessary condition: $f(x)$ is second-order differentiable at x,

$$\bigtriangledown f(x) = 0 \tag{1.8}$$

$$\bigtriangledown^2 f(x) \geq 0 \tag{1.9}$$

1.3.2 Constrained Optimization

Theorem 1.3.1. *Fritz-John Condition*

For constrained optimization problem

$$\begin{aligned} & \min f(x) \\ s.t. \quad & g_i(x) \geq 0, i = 1, ..., n \\ & h_i(x) = 0, i \in 1, ..., m \end{aligned} \tag{1.10}$$

Denote $I(x) = \{i \in \{1, ..., n\} | g_i(x) = 0\}$. *For* $x \in \mathcal{S}$, f *and* $g_i, i \in I(x)$ *is differentiable at* x, $h_j(x)$ *is continuously differentiable at* x. *If* x *is local optimal, then there exists non-trivial* $\lambda_0, \lambda_i \geq 0, i \in I(x)$ *and* μ_j, *such that*

$$\lambda_0 \bigtriangledown f(x) - \sum_{i \in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^{m} \mu_j \bigtriangledown h_j(x) = 0 \tag{1.11}$$

Proof. (i) If $\{\bigtriangledown h_j(x)\}$ is linearly dependent, then there exists non-trivial μ_j, such that

$$\sum_{j=1}^{m} \bigtriangledown \mu_j h_j(x) = 0 \tag{1.12}$$

Let $\lambda_0, \lambda_i, i \in I(x) = 0$, then (1.10) holds.

(ii) If $\{\bigtriangledown h_j(x)\}$ is linearly independent, Denote

$$F_g = F(x, g) = \{d | \bigtriangledown g_i(x)^T d > 0, i \in I(x)\} \tag{1.13}$$

$$F_h = F(x, h) = \{d | \bigtriangledown h_j(x)^T d = 0, j = 1, ..., m\} \tag{1.14}$$

If x is a optimal value, then appearently $F(x,\mathcal{S})\cap D(x,f)=\varnothing$. Due to the independence of $\{\triangledown h_j(x)\}$, we have $F_g\cap F_h\subset F(x,\mathcal{S})$, then

$$F_g\cap F_h\cap D(x,f)=\varnothing \tag{1.15}$$

that is

$$\begin{cases}\triangledown f(x)^T d<0\\ \triangledown g_i(x)^T d>0, i\in I(x)\\ \triangledown h_j(x)^T d=0, j=1,...,m\end{cases} \tag{1.16}$$

has no solution. Let

$$A=\{\triangledown f(x)^T,-\triangledown g_i(x)\}^T, i\in I(x) \tag{1.17}$$

$$B=\{-\triangledown h_j(x)\}, j=1,...,m \tag{1.18}$$

Then (21) is equivalent to

$$\begin{cases}A^T d<0\\ B^T d=0\end{cases} \tag{1.19}$$

has no solution.
Denote

$$S_1=\left\{\begin{pmatrix}y_1\\ y_2\end{pmatrix}|y_1=A^T d, y_2=B^T d, d\in\mathbb{R}^n\right\} \tag{1.20}$$

$$S_2=\left\{\begin{pmatrix}y_1\\ y_2\end{pmatrix}|y_1<0, y_2=0\right\} \tag{1.21}$$

S_1,S_2 are non-trivial convex sets, and $S_1\cap S_2=\varnothing$. From *Hyperplane Separation Theorem*: $\exists\begin{pmatrix}p_1\\ p_2\end{pmatrix}$, such that

$$p_1^T A^T d+p_2^T B^T d\geq p_1^T y_1+p_2^T y_2, \forall d\in\mathbb{R}^n, \forall\begin{pmatrix}y_1\\ y_2\end{pmatrix}\in CL(S_2) \tag{1.22}$$

Let $y_2=0, d=0, y_1<0$, we have

$$p_1\geq 0 \tag{1.23}$$

Let $\begin{pmatrix}y_1\\ y_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}\in CL(S_2)$ So that

$$(p_1^T A^T+p_2^T B^T)d\geq 0 \tag{1.24}$$

$$(Ap_1+Bp_2)^T d\geq 0 \tag{1.25}$$

Let $d = -(Ap_1 + Bp_2)$, we have

$$Ap_1 + Bp_2 = 0 \tag{1.26}$$

From above, we have

$$\begin{cases} Ap_1 + Bp_2 = 0 \\ p_1 \geq 0 \end{cases} \tag{1.27}$$

Let $p_1 = \{\lambda_0, ..., \lambda_{I(x}\}$, $p_2 = \{\mu_1, ..., \mu_m\}$, i.e.,

$$\begin{cases} \lambda_0 \bigtriangledown f(x) - \sum_{i\in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^{m} \mu_j \bigtriangledown h_j(x) = 0 \\ \lambda_i \geq 0 \end{cases} \tag{1.28}$$

□

Theorem 1.3.2. *Kuhn-Tucker Condition*

For constrained optimization problem

$$\begin{array}{ll} & \min f(x) \\ s.t. & g_i(x) \geq 0, i = 1, ..., n \\ & h_i(x) = 0, i \in 1, ..., m \end{array} \tag{1.29}$$

Denote $I(x) = \{i \in \{1, ..., n\} | g_i(x) = 0\}$. *For* $x \in \mathcal{S}$, f *and* $g_i, i \in I(x)$ *is differentiable at* x, $h_j(x)$ *is continuously differentiable at* x. $\{\bigtriangledown g_i(x), i \in I(x); \bigtriangledown h_j(x), j = 1, ..., m\}$ *is linearly independent. If* x *is local optimal, then* $\exists \lambda_i \geq 0$ *and* μ_j*, such that*

$$\bigtriangledown f(x) - \sum_{i\in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^{m} \mu_j \bigtriangledown h_j(x) = 0 \tag{1.30}$$

Remark 1 (K-T condition)**.** *The equation (1.3.2) can be rewritten as*

$$\bigtriangledown f(x) - \sum_{i=1}^{m} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^{m} \mu_j \bigtriangledown h_j(x) = 0 \tag{1.31}$$

where $\lambda_i = 0, i \notin I(x)$. *i.e.,*

$$\lambda_i g_i(x) = 0, i = 1, ..., m \tag{1.32}$$

Denote

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) - \sum_{j=1}^{m} \mu_j h_j(x) \tag{1.33}$$

as the Lagrange function*, then the K-T condition can be formulated as*

$$(K-T)\begin{cases} \bigtriangledown_{\mathbf{x}}\mathcal{L}(\mathbf{x},\lambda,\mu)=0 \\ \bigtriangledown_{\lambda}\mathcal{L}(\mathbf{x},\lambda,\mu)=0 \\ \bigtriangledown_{\mu}\mathcal{L}(\mathbf{x},\lambda,\mu)=0 \\ \lambda_i \geq 0, i=1,...,m \\ \lambda_i g_i(\mathbf{x})=0, i=1,..,m \end{cases} \tag{1.34}$$

1.4 Descent function

Definition 1.4.1. *Descent function. Denote solution set $\Omega \in X$, $\mathcal{A}$ is an algorithm on X, $\psi : X \to \mathbb{R}$. If*

$$\psi(y) < \psi(x), \quad \forall x \notin \Omega, y \in \mathcal{A}(x) \tag{1.35}$$

$$\psi(y) \leq \psi(x), \quad \forall x \in \Omega, y \in \mathcal{A}(x) \tag{1.36}$$

Then ψ is a ***descent function*** *of* $(\Omega, \mathcal{A})$.

1.5 Convergence of Algorithm

Theorem 1.5.1. *$\mathcal{A}$ is an algorithm on X, Ω is the solution set, $x^{(0)} \in X$. If $x^{(k)} \in \Omega$, then the iteration stops. Otherwise set $x^{(k+1)} = \mathcal{A}(x^{(k)}), k := k+1$. If*

- *$\{x^{(k)}\}$ in a compact subset of X*
- *There exists a continuous function ψ, ψ is a descent function of* $(\Omega, \mathcal{A})$
- *$\mathcal{A}$ is closed on Ω^C*

Then, any convergent subsequence of $\{x^{(k)}\}$ converges to $x, x \in \Omega$.

Proof. □

1.5.1 Search Methods

1.5.1.1 Line Search

Generate $d^{(k)}$ from $x^{(k)}$,

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \tag{1.37}$$

. search α_k in 1-D space.

1.5.1.2 Trust Region

Generate local model $Q_k(s)$ of $x^{(k)}$,

$$s^{(k)} = \arg\min Q_k(s) \tag{1.38}$$

$$x^{(k+1)} = x^{(k)} + s^{(k)} \tag{1.39}$$

2 Unconstrained Optimization

2.1 Gradient Based Methods

$$\min_{x \in \mathbb{R}^n} f(x) \tag{2.1}$$

Algorithm 1: Example of gradient based algorithm

Data: Solution set Ω, cost function f
$x^{(0)} \in \mathbb{R}^n, k := 0$;
while $x^{(k)} \notin \Omega$ **do**
  $d^{(k)} = -H_k \bigtriangledown f(x^{(k)})$, ($H_k$ is a positive definite symmetrical matrix);
  solve $\min_{\alpha_k \geq 0} f(x^{(k)} + \alpha_k d^{(k)})$;
  $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$, $k := k + 1$
end

2.1.1 Determine Search Direction

2.1.1.1 First-order gradient method

For unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} f(x) \tag{2.2}$$

We have

$$f(x) = f(x^{(k)}) + \bigtriangledown f(x^{(k)})^T (x - x^{(k)}) + O(\| x - x^{(k)} \|^2) \tag{2.3}$$

Set $d^{(k)} = - \bigtriangledown f(x^{(k)})$, when α_k is sufficiently small,

$$f(x^{(k)} + \alpha_k d^{(k)}) < f(x^{(k)}) \tag{2.4}$$

2.1.1.2 Second-order gradient method – Newton Direction

$$\begin{aligned} f(x) =& f(x^{(k)}) + \bigtriangledown f(x^{(k)})^T (x - x^{(k)}) & (2.5)\\ &+ \frac{1}{2}(x - x^{(k)})^T \bigtriangledown^2 f(x^{(k)})(x - x^{(k)}) + O(\| x - x^{(k)} \|^3) & (2.6) \end{aligned}$$

Set $d^{(k)} = -G_k^{-1} \bigtriangledown f(x^{(k)})$, where $G_k = \bigtriangledown^2 f(x^{(k)})$, i.e., Hesse matrix of f at $x^{(k)}$.

2.1.2 Determine Step Factor – Line Search

$$\min_{\alpha \geq 0} \varphi(\alpha) = f(x^{(k)} + \alpha d^{(k)}) \tag{2.7}$$

2.1.2.1 Exact Line Search

Solve Line Search problem in finite iterations.

2.1.2.2 Inexact Line Search

In some cases, the exact solution of Line Search is not necessary, so we can use inexace line search to improve algorithm efficiency.

Goldstein Conditions

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{2.8}$$

$$\varphi(\alpha) \geq \varphi(0) + (1-\rho)\alpha\varphi'(0) \tag{2.9}$$

where$\rho \in (\frac{1}{2}, 1)$ is a fixed parameter.

However, the downside of Goldstein Conditions is that the optimal value might not lie in the valid area.

Wolfe-Powell Conditions

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{2.10}$$

$$\varphi'(\alpha) \geq \sigma\varphi'(0) \tag{2.11}$$

where $\sigma \in (\rho, 1)$.

2.1.3 Global Convergence

Theorem 2.1.1. *Assume f continuously differentiable on level set $L(x^{(0)}) = \{x|f(x) \leq f(x^{(0)})\}$. Denote $\theta^{(k)}$ as the angle between $d^{(k)}$ and $- \bigtriangledown f(x^{(k)})$.*

$$\theta^{(k)} \leq \frac{\pi}{2} - \mu \tag{2.12}$$

If step factor is determined by following methods

- *Exace Line Search*
- *Goldstein Conditions*
- *Wolfe-Powell Conditions*

Then, there exists k, such that $\bigtriangledown f(x^{(k)}) = 0$, or $f(x^{(k)}) \to 0$ or $f(x^{(k)}) \to -\infty$.

Proof. (In the Wolfe-Powell Conditions case)

Suppose for all $k, \mathbf{g}^{(k)} = \bigtriangledown f(\mathbf{x}^{(k)}) \neq 0$ and $f(\mathbf{x}^{(k)})$ has finite lower bound. From (2.12), we have $\mathbf{d}^{(k)}$ is descent direction at point $\mathbf{x}^{(k)}$. So from Wolfe-Powell conditions, $f(\mathbf{x}^{(k)})$

decrease monotonically, so $f(\mathbf{x}^{(k)})$ is convergent sequence, then

$$f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k+1)}) \to 0 \tag{2.13}$$

From (2.10), we have

$$-\rho\alpha\varphi'(0) \leq \varphi(0) - \varphi(\alpha) \tag{2.14}$$
$$-\rho\alpha\mathbf{g}^{(k)T}\mathbf{d}^{(k)} \leq f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k+1)}) \tag{2.15}$$
$$-\mathbf{g}^{(k)T}\mathbf{s}^{(k)} \leq \frac{f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k+1)})}{\rho} \tag{2.16}$$

i.e.,

$$-\mathbf{g}^{(k)T}\mathbf{s}^{(k)} \to 0 \tag{2.17}$$

If $\mathbf{g}^{(k)} \to 0$ do not hold, i.e., $\exists\varepsilon > 0$ and subsequence $\{\mathbf{x}^{(k)}\}$ such that $\| \mathbf{g}^{(k)} \| \geq \varepsilon$, so

$$-\mathbf{g}^{(k)T}\mathbf{s}^{(k)} = \| \mathbf{g}^{(k)} \| \| \mathbf{s}^{(k)} \| \cos\theta_k \geq \varepsilon \| \mathbf{s}^{(k)} \| \sin\mu \tag{2.18}$$

then

$$\| \mathbf{s}^{(k)} \| \to 0 \tag{2.19}$$

Due to the continuously differentiability of f,

$$\mathbf{g}^{(k+1)T}\mathbf{s}^{(k)} - \mathbf{g}^{(k)T}\mathbf{s}^{(k)} = (\bigtriangledown f(\mathbf{x}^{(k+1)}) - \bigtriangledown f(\mathbf{x}^{(k)}))^T\mathbf{s}^{(k)} \tag{2.20}$$
$$= (\bigtriangledown^2 f(\mathbf{x}^{(k)})\mathbf{s}^{(k)} + o(\mathbf{s}^{(k)}))^T\mathbf{s}^{(k)} \tag{2.21}$$
$$= \mathbf{s}^{(k)T} \bigtriangledown^2 f(\mathbf{x}^{(k)})\mathbf{s}^{(k)} + o(\mathbf{s}^{(k)}))^T\mathbf{s}^{(k)} \tag{2.22}$$
$$= o(\| \mathbf{s}^{(k)} \|) \tag{2.23}$$

then

$$\frac{\mathbf{g}^{(k+1)T}\mathbf{s}^{(k)}}{\mathbf{g}^{(k)T}\mathbf{s}^{(k)}} \to 1 \tag{2.24}$$

is conflict with (2.11), so

$$\mathbf{g}^{(k)} \to 0 \tag{2.25}$$

□

2.1.4 Steepest Descent Method

Steepest Descent Method is a Line Search Method.

$$x^{(k+1)} = x^{(k)} - \alpha_k \bigtriangledown f(x^{(k)}) \tag{2.26}$$

Algorithm 2: Steepest Descent Algorithm

Data: Termination error ϵ, cost function f

$x^{(0)} \in \mathbb{R}^n, k := 0$;

while $\| g^{(k)} \| \geq \epsilon$ **do**

  $d^{(k)} = -g^{(k)}$;

  solve $\min_{\alpha_k \geq 0} f(x^{(k)} + \alpha_k d^{(k)})$;

  $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$, $k := k + 1$;

  Compute $g^{(k)} = \bigtriangledown f(x^{(k)})$

end

Steepest Descent Method has linear convergence rate generally.

2.1.5 Newton Method

Newton Method is also a Line Search Method.

$$f(x^{(k)} + s) \approx q^{(k)}(s) f(x^{(k)}) + g^{(k)^T} s + \frac{1}{2} s^T G_k s \tag{2.27}$$

where $g^{(k)} = \bigtriangledown f(x^{(k)})$, $G_k = \bigtriangledown^2 f(x^{(k)})$. To minimize $q^{(k)}(s)$, we have

$$s = G_k^{-1} g^{(k)} \tag{2.28}$$

Notice that $G_k^{-1} g^{(k)}$ is the Newton Direction.

Analysis on quadratic function

For positive definite quadratic function

$$f(x) = \frac{1}{2} x^T G x - c^T x \tag{2.29}$$

In this case, $\bigtriangledown^2 f(x) = G$. Let $H_0 = G^{-1}$, then we have

$$d^{(0)} = H_0 \bigtriangledown f(x^{(0)}) \tag{2.30}$$
$$= G^{-1}(Gx^{(0)} - c) \tag{2.31}$$
$$= x^{(0)} - G^{-1} c \tag{2.32}$$
$$= x^{(0)} - x^* \tag{2.33}$$

So that Newton Method can reach global optimal in 1 iteration for quadratic functions.

For general non-linear functions, if we follow

$$x^{(k+1)} = x^{(k)} - G_k^{-1} g^{(k)} \tag{2.34}$$

we called it Newton Method.

Convergence Rate of Newton Method

Theorem 2.1.2. *$f \in \mathcal{C}^2$, $x^{(k)}$ is sufficiently closed to optimal point x^*, where $\bigtriangledown f(x^*) = 0$. If $\bigtriangledown^2 f(x^*)$ is positive definite, Hesse matrix of f satisfies Lipschitz Condition, i.e., $\exists \beta > 0$, such that for all (i, j),*

$$|G_{ij}(x) - G_{ij}(y)| \leq \beta \parallel x - y \parallel \tag{2.35}$$

Then $\{x^{(k)}\} \to x^*$, and have quadratic convergence rate.

Proof. Denote $g(x) = \bigtriangledown f(x)$, then we have

$$g(x - h) = g(x) - G(x)h + O(\parallel h \parallel^2) \tag{2.36}$$

Let $x = x^{(k)}$, $h = h^{(k)} = x^{(k)} - x^*$, then

$$g(x^*) = g(x^{(k)}) - G(x^{(k)})(h^{(k)}) + O(\parallel h^{(k)} \parallel^2) = 0 \tag{2.37}$$

From Lipschitz Condition, we can easily get $G(x^{(k)})^{-1}$ is finite. Then we left multiply $G(x^{(k)})^{-1}$ to Equation (2.37)

$$\begin{aligned} 0 &= G(x^{(k)})^{-1} g(x^{(k)}) - h^{(k)} + O(\parallel h^{(k)} \parallel^2) && (2.38)\\ &= x^* - x^{(k)} + G(x^{(k)})^{-1} g(x^{(k)}) + O(\parallel h^{(k)} \parallel^2) && (2.39)\\ &= x^* - x^{(k+1)} + O(\parallel h^{(k)} \parallel^2) && (2.40)\\ &= -h^{(k+1)} + O(\parallel h^{(k)} \parallel^2) && (2.41) \end{aligned}$$

i.e.,

$$\parallel h^{(k+1)} \parallel = O(\parallel h^{(k)} \parallel^2) \tag{2.42}$$

□

2.1.6 Quasi-Newton Method

Newton Method has a fast convergence rate. However, Newton Method requires second-order derivative, if Hesse matrix is not positive definite, Newton Method might not work well.

In order to overcome the above difficulties, Quasi-Newton Method is introduced. Its basic idea is that: Using second-order derivative free matrix H_k to approximate $G(x^{(k)})^{-1}$. Denote $s^{(k)} = x^{(k+1)} - x^{(k)}$, $y^{(k)} = \bigtriangledown f(x^{(k+1)}) - \bigtriangledown f(x^{(k)})$, then we have

$$\bigtriangledown^2 f(x^{(k)}) s^{(k)} \approx y^{(k)} \tag{2.43}$$

or

$$\bigtriangledown^2 f(x^{(k)})^{-1} y^{(k)} \approx s^{(k)} \tag{2.44}$$

So we need to construct H_{k+1} such that

$$H_{k+1}y^{(k)} \approx s^{(k)} \tag{2.45}$$

or

$$y^{(k)} \approx B_{k+1}s^{(k)} \tag{2.46}$$

we called (2.45), (2.46) *Quasi-Newton Conditions* or *Secant Conditions*.

Algorithm 3: Quasi-Newton Algorithm

Data: Cost function f
$x^{(0)} \in \mathbb{R}^n, H_0 = I, k := 0;$
while *some conditions* **do**
  $d^{(k)} = -H_k g^{(k)};$
  solve $\min_{\alpha_k \geq 0} f(x^{(k)} + \alpha_k d^{(k)});$
  $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)};$
  generate H_{k+1}, $k := k + 1$
end

2.1.6.1 How to generate H_k

H_k is the approximation matrix in kth iteration, we want to generate H_{k+1} from H_k

Symmetric Rank 1 Update Assume

$$H_{k+1} = H_k + a\mathbf{u}\mathbf{u}^T, \quad a \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^n \tag{2.47}$$

From the Quasi-Newton Conditions, we have

$$H_{k+1}\mathbf{y}^{(k)} = \mathbf{s}^{(k)} \tag{2.48}$$

$$H_k\mathbf{y}^{(k)} + a\mathbf{u}\mathbf{u}^T\mathbf{y}^{(k)} = \mathbf{s}^{(k)} \tag{2.49}$$

$$H_k\mathbf{y}^{(k)} + a\mathbf{u}^T\mathbf{y}^{(k)}\mathbf{u} = \mathbf{s}^{(k)} \tag{2.50}$$

Let $\mathbf{u} = \mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)}$, $a = \frac{1}{\mathbf{u}^T\mathbf{y}}$, clearly this is a solution of the equation. Here we have

$$H_{k+1} = \frac{(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})^T}{(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})^T\mathbf{y}^{(k)}} \tag{2.51}$$

(2.51) is *Symmetric Rank 1 Update*. The problem of Symmetric Rank 1 Update is that the positive-definite property of H_k can not be preserved.

Symmetric Rank 2 Update Assume

$$H_{k+1} = H_k + a\mathbf{u}\mathbf{u}^T + b\mathbf{v}\mathbf{v}^T, \quad a, b \in \mathbb{R}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \tag{2.52}$$

such that Quasi-Newton Conditions stand. We can find a solution of $a, b, \mathbf{u}, \mathbf{v}$ that is

$$\begin{cases} \mathbf{u} = \mathbf{s}^{(k)}, \quad a\mathbf{u}^T\mathbf{y} = 1 \\ \mathbf{v} = H_k\mathbf{y}^{(k)}, \quad b\mathbf{v}^T\mathbf{y} = -1 \end{cases} \tag{2.53}$$

So that we have

$$H_{k+1} = H_k + \frac{\mathbf{s}^{(k)}\mathbf{s}^{(k)T}}{\mathbf{s}^{(k)T}\mathbf{y}^{(k)}} - \frac{H_k\mathbf{y}^{(k)}\mathbf{y}^{(k)T}H_k}{\mathbf{y}^{(k)T}H_k\mathbf{y}^{(k)}} \tag{2.54}$$

We called (2.54) the DFP (Davidon-Fletcher-Powell) update.

From Quasi-Newton Condition (2.46), we can get the BFGS (Broyden-Fletcher-Goldfarb-Shanno) update

$$B_{k+1}^{(BFGS)} = B_k + \frac{\mathbf{y}^{(k)}\mathbf{y}^{(k)T}}{\mathbf{y}^{(k)T}\mathbf{s}^{(k)}} - \frac{B_k\mathbf{s}^{(k)}\mathbf{s}^{(k)T}B_k}{\mathbf{s}^{(k)T}B_k\mathbf{s}^{(k)}} \tag{2.55}$$

Inverse of SR1 update

Theorem 2.1.3 (Sherman-Morrison). *$A \in \mathbb{R}^n \times \mathbb{R}^n$ is a non-singular matrix, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If $1 + \mathbf{v}^T A^{-1}\mathbf{u} \neq 0$, then SR1 update of A is non-singular, and its inverse can be represented as*

$$(A + a\mathbf{u}\mathbf{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^T A^{-1}}{1 + \mathbf{v}^T A^{-1}\mathbf{u}} \tag{2.56}$$

2.1.7 Conjugate Gradient Method

Definition 2.1.1. *Conjugate Direction. G is a $n \times n$ positive definite matrix, for non-zero vector set $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\} \in \mathbb{R}^n$, if $\mathbf{d}^{(i)T}G\mathbf{d}^{(j)} = 0, (i \neq j)$, then we called $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ is* G-Conjugate.

Lemma 2.1.4. *For non-zero conjugate vector set $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\} \in \mathbb{R}^n$, $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ are linearly independent.*

Proof. From Definition 2.1.1, we have

$$\mathbf{d}^{(i)T}G\mathbf{d}^{(j)} = 0, \forall i, j, i \neq j \tag{2.57}$$

if $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ is linearly dependent, there exists

$$\mathbf{d}^{(t)} = \sum_{j=0}^{k} c_j\mathbf{d}^{(j)} \tag{2.58}$$

then

$$\mathbf{d}^{(t)T}G\mathbf{d}^{(i)} = \sum_{j=0}^{k} c_j\mathbf{d}^{(j)}G\mathbf{d}^{(i)} = c_i\mathbf{d}^{(i)}G\mathbf{d}^{(i)} \neq 0 \tag{2.59}$$

so that $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ are linearly independent.

$\square$

Algorithm 4: Conjuagte Gradient Algorithm

Data: Cost function f

$x^{(0)} \in \mathbb{R}^n$, positive definite matrix G, $k := 0$;

Construct $\mathbf{d}^{(0)}$ such that $\mathbf{g}^{(0)T}\mathbf{d}^{(0)} < 0$;

while *some conditions* **do**

  solve $\min_{\alpha_k \geq 0} f(\mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)})$;

  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$;

  Construct $\mathbf{d}^{(k+1)}$ such that $\mathbf{d}^{(k+1)} G \mathbf{d}^{(j)} = 0, j = 0, ..., k.$;

  $k := k + 1$

end

Theorem 2.1.5 (Conjugate Gradient)**.** *For strictly convex quadratic function* $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G \mathbf{x} + \mathbf{c}^T \mathbf{x}$*, apply conjugate gradient method combined with exact line search, then* $\mathbf{x}^{(k+1)}$ *is the global minima in manifold*

$$\mathcal{V} = \{\mathbf{x} | \mathbf{x} = \mathbf{x}^{(0)} + \sum_{j=0}^{k} \beta_j \mathbf{d}^{(j)}, \forall \beta_j \in \mathbb{R}\} \tag{2.60}$$

Proof. Firstly, from Lemma 2.1.6, we have $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ are linearly independent. So we only need to prove that for all $k < n$

$$\mathbf{g}^{(k+1)T}\mathbf{d}^{(j)} = 0, j = 0, ..., k \tag{2.61}$$

i.e., $\mathbf{g}^{(k+1)}$ is orthogonal with subspace $span\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$.

Due to the exact line search, $\forall j$

$$\mathbf{g}^{(j+1)T}\mathbf{d}^{(j)} = 0 \tag{2.62}$$

especially $\mathbf{g}^{(k+1)T}\mathbf{d}^{(k)} = 0$.

Notice that

$$\mathbf{g}^{(k+1)} - \mathbf{g}^{(k)} = G(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \alpha_k G \mathbf{d}^{(k)} \tag{2.63}$$

so that we have $\forall j \leq k$

$$\mathbf{g}^{(k+1)T}\mathbf{d}^{(j)} = (\sum_{m=j+1}^{k}(\mathbf{g}^{(m+1)T} - \mathbf{g}^{(m)T}) + \mathbf{g}^{(j+1)T})\mathbf{d}^{(j)} \tag{2.64}$$

$$= \sum_{m=j+1} \alpha_m \mathbf{d}^{(m)T} G \mathbf{d}^{(j)} + \mathbf{g}^{(j+1)T}\mathbf{d}^{(j)} \tag{2.65}$$

$$= 0 \tag{2.66}$$

□

Lemma 2.1.6. *For strictly convex quadratic function* $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x}$*, apply conjugate gradient method combined with exact line search,* $\mathbf{g}(\mathbf{x}) = \bigtriangledown f(\mathbf{x}) = G\mathbf{x} + \mathbf{c}$*, we have*

$$\mathbf{g}^{(k)T}\mathbf{g}^{(j)} = 0, \forall j = 0, ..., k-1 \tag{2.67}$$

Proof. From Theorem 2.1.5, we have

$$\mathbf{g}^{(k)T}\mathbf{g}^{(j)} = \mathbf{g}^{(k)T}(-\mathbf{d}^{(j)} + \sum_{i=0}^{j-1}\beta_i^{(j)}\mathbf{d}^{(i)}) = 0 \tag{2.68}$$

□

2.1.7.1 Quadratic function case

For $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T x$, G is a $n \times n$ positive definite matrix.

$$\mathbf{g}(\mathbf{x}) = G\mathbf{x} + \mathbf{c} \tag{2.69}$$

Set $\mathbf{d}^{(0)} = -\mathbf{g}^{(0)}$, exact line search for α_0 such that $\mathbf{g}^{(1)T}\mathbf{d}^{(0)} = 0$. Assume $\mathbf{d}^{(1)} = -\mathbf{g}^{(1)} + \beta_0^{(1)}\mathbf{d}^{(0)}$, select $\beta_0^{(1)}$ such that $\mathbf{d}^{(1)}G\mathbf{d}^{(0)} = 0$

$$\beta_0^{(1)} = \frac{\mathbf{g}^{(1)T}\mathbf{g}^{(1)}}{\mathbf{g}^{(0)T}\mathbf{g}^{(0)}} \tag{2.70}$$

Proof. From (92), we have

$$\mathbf{d}^{(1)T}G\mathbf{d}^{(0)} = 0 \tag{2.71}$$

$$\Leftrightarrow \quad \mathbf{d}^{(1)T}(\mathbf{g}^{(1)} - \mathbf{g}^{(0)}) = 0 \tag{2.72}$$

$$\Leftrightarrow \quad (\mathbf{g}^{(1)} + \beta_0^{(1)}\mathbf{g}^{(0)})^T(\mathbf{g}^{(1)} - \mathbf{g}^{(0)}) = 0 \tag{2.73}$$

$$\Leftrightarrow \quad \mathbf{g}^{(1)T}\mathbf{g}^{(1)} - \beta_0^{(1)}\mathbf{g}^{(0)T}\mathbf{g}^{(0)} = 0 \tag{2.74}$$

$$\Leftrightarrow \quad \beta_0^{(1)} = \frac{\mathbf{g}^{(1)T}\mathbf{g}^{(1)}}{\mathbf{g}^{(0)T}\mathbf{g}^{(0)}} \tag{2.75}$$

□

Generally, we can select $\beta_j^{(k)}$ such that $\mathbf{d}^{(k)T}G\mathbf{d}^{(j)}=0, j=0,1,...,k-1$ that is

$$\mathbf{d}^{(k)T}G\mathbf{d}^{(j)}=0 \tag{2.76}$$

$$(-\mathbf{g}^{(k)T}+\sum_{i=0}^{k-1}\beta_i^{(k)}\mathbf{d}^{(i)T})G\mathbf{d}^{(j)}=0 \tag{2.77}$$

$$-\mathbf{g}^{(k)T}G\mathbf{d}^{(j)}+\beta_j^{(k)}\mathbf{d}^{(j)T}G\mathbf{d}^{(j)}=0 \tag{2.78}$$

so we have

$$\beta_j^{(k)}=\frac{\mathbf{g}^{(k)T}G\mathbf{d}^{(j)}}{\mathbf{d}^{(j)T}G\mathbf{d}^{(j)}}=\frac{\mathbf{g}^{(k)T}(\mathbf{g}^{(j+1)}-\mathbf{g}^{(j)})}{\mathbf{d}^{(j)T}(\mathbf{g}^{(j+1)}-\mathbf{g}^{(j)})} \tag{2.79}$$

From Lemma 2.1.6, we have

$$\mathbf{g}^{(k)T}\mathbf{g}^{(j)}=0, \forall j=0,...,k-1 \tag{2.80}$$

So

$$\beta_j^{(k)}=0, j=0,...,k-2 \tag{2.81}$$

$$\beta_{k-1}^{(k)}=\frac{\mathbf{g}^{(k)T}(\mathbf{g}^{(k)}}{\mathbf{g}^{(k-1)T}(\mathbf{g}^{(k-1)}} \tag{2.82}$$

2.2 Trust Region Method

Previously, we use a direction search strategy to determine a search direction, then use line search method to determine step length.

Now we discuss a new global convergence strategy – Trust-Region Method.

Definition 2.2.1 (Trust Region)**.**

$$\Omega_k=\{\mathbf{x}\in\mathbb{R}^n \mid \ \| \mathbf{x}-\mathbf{x}^{(k)} \| \le e_k\} \tag{2.83}$$

We called Ω_k Trust Region*,* e_k *is the* Trust radius*.*

Suppose in this neighborhood, quadratic model $q^{(k)}(\mathbf{s})$ is a proper approximation of $f(\mathbf{x})$. We minimize the quadratic model in trust region, derive approximate minima $\mathbf{s}^{(k)}$, and set $\mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}+\mathbf{s}^{(k)}$.

2.2.1 Trust Region Subproblem

$$\min_{\|\mathbf{s}\|\le e_k} q^{(k)}(\mathbf{s})=f(\mathbf{x}^{(k)})+\mathbf{g}^{(k)T}\mathbf{s}+\frac{1}{2}\mathbf{s}^TB_k\mathbf{s} \tag{2.84}$$

Where $\mathbf{s}=\mathbf{x}-\mathbf{x}^{(k)}$, $\mathbf{g}^{(k)}=\bigtriangledown f(\mathbf{x}^{(k)})$, $B_k=\bigtriangledown^2 f(\mathbf{x}^{(k)})$. e_k is the trust region radius.

2.2.2 How to select e_k

Denote the solution of the subproblem as $\mathbf{s}^{(k)}$, then let

$$\text{Act}_k = f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k)} + \mathbf{s}^{(k)}) \tag{2.85}$$

$$\text{Pre}_k = q^{(k)}(\mathbf{0}) - q^{(k)}(\mathbf{s}^{(k)}) \tag{2.86}$$

Define

$$r_k = \frac{\text{Act}_k}{\text{Pre}_k} = \frac{f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k)} + \mathbf{s}^{(k)})}{q^{(k)}(\mathbf{0}) - q^{(k)}(\mathbf{s}^{(k)})} \tag{2.87}$$

to measure the difference between objective function and the quadratic approximate model.

We can update e_k according to r_k. If r_k is too small, that means our model can not fit the objective function well, so we need to decrease e_k. If r_k is close to 1, that means out model is good and we can increase r_k. Set the parameters $0 < \gamma_1 < \gamma_2 < 1$ and $0 < \eta_1 < 1 < \eta_2$, we can have the following update rule

$$e_{k+1} = \begin{cases} \eta_1 e_k & \text{if } r_k < \gamma_1 \\ e_k & \text{if } \gamma_1 < r_k < \gamma_2 \\ \min(\eta_2 e_k, \bar{e}) & \text{if } r_k \geq \gamma_2 \end{cases} \tag{2.88}$$

Algorithm 5: Trust Region Algorithm

Data: Cost function f
$x^{(0)} \in \mathbb{R}^n$, $e_0 \in (0, \bar{e})$, $\epsilon > 0$, $0 < \gamma_1 < \gamma_2 < 1$, $0 < \eta_1 < 1 < \eta_2$, $k := 0$;

while $\| \mathbf{g}^{(k)} \| \geq \epsilon$ **do**
  solve the subproblem to derive $\mathbf{s}^{(k)}$;
  calculate r_k, update $\mathbf{x}$;
  if $r_k > 0$ **then**
    $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$
  else
    $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$
  end
  update e_k following (117);
  $k := k + 1$;
end

3 Constrained Optimization

3.1 Quadratic Programming

$$
\begin{array}{lll}
\min & Q(\mathbf{x}) & = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \\
s.t. & \mathbf{a}_i^T\mathbf{x} & = b_i, i \in \mathcal{E} = \{1, ..., m_e\} \\
 & \mathbf{a}_i^T\mathbf{x} & \geq b_i, i \in \mathcal{I} = \{m_e + 1, ..., m\}
\end{array}
\tag{3.1}
$$

We assume that G is a symmetric matrix and $\mathbf{a}_i, i \in \mathcal{E}$ be linearly independent.

3.1.1 Solution of Quadratic Programming

If G be positive semi-definite matrix, the Quadratic Programming problem is a convex optimization problem, so any of its local minima is a global minima.

If G be positive definite matrix, the solution to the Quadratic Programming problem is unique, if exists.

If G be indefinite, there is no guarantee to the solution.

3.1.2 Equality Constrained Quadratic Programming

$$
\begin{array}{ll}
\min\ Q(\mathbf{x}) & = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \\
s.t.\ A\mathbf{x} & = \mathbf{b}
\end{array}
\tag{3.2}
$$

3.1.3 General Quadratic Programming

$$
\begin{array}{lll}
\min & Q(\mathbf{x}) & = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \\
s.t. & \mathbf{a}_i^T\mathbf{x} & = b_i, i \in \mathcal{E} = \{1, ..., m_e\} \\
 & \mathbf{a}_i^T\mathbf{x} & \geq b_i, i \in \mathcal{I} = \{m_e + 1, ..., m\}
\end{array}
\tag{3.3}
$$

The idea is to remove or transform the inequality constraints. If the inequality constraint is not active near the solution, we can ignore the constraint; For the active inequality constraints, we can use equality constraints to replace them.

Theorem 3.1.1 (Active Set)**.** *Denote $\mathbf{x}^*$ as a local minima of general quadratic problem* (3.3)*, then $\mathbf{x}^*$ must be a local minima of the equality constrained problem*

$$
\text{(EQ)} \left\{
\begin{array}{ll}
\min\ Q(\mathbf{x}) & = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \\
s.t.\ \mathbf{a}_i^T\mathbf{x} & = b_i, i \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)
\end{array}
\right.
\tag{3.4}
$$

Meanwhile, if $\mathbf{x}^*$ *is a feasible point of* (3.3)*, and the K-T point of (EQ),* $\lambda^* \geq 0, i \in \mathcal{I}(\mathbf{x}^*)$*, then* $\mathbf{x}^*$ *must be the K-T point of* (3.3).

Proof. Recall the K-T condition, we can get that there exists $\lambda_i \geq 0, i \in I(\mathbf{x}^*)$ and μ_j s.t.

$$\triangledown Q(\mathbf{x}^*) - \sum_{i \in I(\mathbf{x}^*)} \lambda_i \mathbf{a}_i - \sum_{j \in \mathcal{E}} \mu_j \mathbf{a}_j = 0 \tag{3.5}$$

the K-T condition of (EQ) is there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)$, s.t.

$$\triangledown Q(\mathbf{x}^*) - \sum_{j \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)} \lambda_j \mathbf{a}_j = 0 \tag{3.6}$$

Appearently If $\mathbf{x}^*$ satisfies (3.5), then it also satisfies (3.6). On the other hand, if $\mathbf{x}^*$ satisfies (3.6) and $\lambda_i \geq 0, i \in I(\mathbf{x}^*)$, we have

$$\triangledown\, Q(\mathbf{x}^*) - \sum_{j \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)} \lambda_j \mathbf{a}_j = 0 \tag{3.7}$$

$$\Leftrightarrow\ \triangledown\, Q(\mathbf{x}^*) - \sum_{i \in I(\mathbf{x}^*)} \lambda_i \mathbf{a}_i - \sum_{j \in \mathcal{E}} \lambda_j \mathbf{a}_j = 0 \tag{3.8}$$

i.e., $\mathbf{x}^*$ satisfies (3.5).

□

3.2 Equality Constrained Problem

3.2.1 Lagrange-Newton method

$$\min f(\mathbf{x}) \tag{3.9}$$

$$s.t.\ \mathbf{c}(\mathbf{x}) = \mathbf{0} \tag{3.10}$$

where $\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), ..., c_m(\mathbf{x}))^T$.

Denote $A(\mathbf{x}) = [\triangledown \mathbf{c}(\mathbf{x})]^T = (\triangledown c_1(\mathbf{x}), ..., \triangledown c_m(\mathbf{x}))^T$. The K-T condition of the problem is there exists $\lambda \in \mathbb{R}^m$ s.t.

$$\triangledown f(\mathbf{x}) - A(\mathbf{x})^T \lambda = \mathbf{0} \tag{3.11}$$

and $\mathbf{c}(\mathbf{x}) = \mathbf{0}$.

We can use Newton-Raphson method to solve the equations by

$$\begin{pmatrix} W(\mathbf{x}, \lambda) & -A(\mathbf{x})^T \\ -A(\mathbf{x}) & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = - \begin{pmatrix} \triangledown f(\mathbf{x}) - A(\mathbf{x})^T \lambda \\ \mathbf{c}(\mathbf{x}) \end{pmatrix} \tag{3.12}$$

where $W(\mathbf{x}, \lambda) = \triangledown^2 f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \triangledown^2 c_i(\mathbf{x})$.

We called the method above as *Lagrange-Newton Method*.

Here we can define

$$\psi(\mathbf{x},\lambda) = \parallel \bigtriangledown f(\mathbf{x}) - A(\mathbf{x})^T\lambda \parallel^2 + \parallel \mathbf{c}(\mathbf{x}) \parallel^2 \tag{3.13}$$

so that ψ is a descent function to Lagrange-Newton method.

$$\bigtriangledown\psi(\mathbf{x},\lambda)^T \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = -2\psi(\mathbf{x},\lambda) \neq 0 \tag{3.14}$$

3.2.2 Sequential Quadratic Programming method

(3.12) can be rewritten into

$$\begin{cases} W(\mathbf{x},\lambda)\delta_x + \bigtriangledown f(\mathbf{x}) & = A(\mathbf{x})^T(\lambda + \delta_\lambda) \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\delta_x & = \mathbf{0} \end{cases} \tag{3.15}$$

From K-T condition, we notice that δ_x is the K-T point of the following Quadratic Programming problem

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{d}^T W(\mathbf{x},\lambda)\mathbf{d} + \bigtriangledown f(\mathbf{x})^T\mathbf{d} \\ s.t. & \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} = 0 \end{array} \tag{3.16}$$

So we can solve a Quadratic Programming subproblem to derive δ_x, we called this method *Sequential Quadratic Programming*.

3.3 General Nonlinear Constrained Problem

3.3.1 Sequential Quadratic Programming method

$$\begin{array}{lll} \min & f(\mathbf{x}) & \\ s.t. & c_i(\mathbf{x}) = 0, & i \in \mathcal{E} = \{1, ..., m_e\} \\ & c_i(\mathbf{x}) \geq 0, & i \in \mathcal{I} = \{m_e + 1, ..., m\} \end{array} \tag{3.17}$$

Similarly, we can construct subproblem

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{d}^T W\mathbf{d} + \mathbf{g}^T\mathbf{d} \\ s.t. & c_i(\mathbf{x}) + \mathbf{a}_i(\mathbf{x})^T\mathbf{d} = 0, i \in \mathcal{E} \\ & c_i(\mathbf{x}) + \mathbf{a}_i(\mathbf{x})^T\mathbf{d} \geq 0, i \in \mathcal{I} \end{array} \tag{3.18}$$

Here, W is the Hesse matrix (or its approximation) of the Lagrange function of (3.17), $\mathbf{g} = \bigtriangledown f(\mathbf{x})$, $A(\mathbf{x}) = (\mathbf{a}_1(\mathbf{x}), ..., \mathbf{a}_m(\mathbf{x})$.

Denote the solution to subproblem (3.18) as $\mathbf{d}$, the corresponding Lagrange multiplier vector $\bar{\lambda}$, so we have

$$\begin{cases} W\mathbf{d} + \mathbf{g} = A(\mathbf{x})^T\bar{\lambda} \\ \bar{\lambda}_i \geq 0, i \in \mathcal{I} \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} = 0, i \in \mathcal{E} \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} \geq 0, i \in \mathcal{I} \end{cases} \tag{3.19}$$

3.3.2 Penalty method

For nonlinear constrained porblem (3.17), we can use objective function $f(\mathbf{x})$ and constraint function $\mathbf{c}(\mathbf{x})$ to construct *Penalty function*

$$P(\mathbf{x}) = P(f(\mathbf{x}), \mathbf{c}(\mathbf{x})) \tag{3.20}$$

We need the penalty function have the property that: for feasible points, $P(\mathbf{x}) = f(\mathbf{x})$, otherwise, $P(\mathbf{x}) > f(\mathbf{x})$.

To measure the destructiveness to the constraints, we define $\mathbf{c}(\mathbf{x})_-$

$$\begin{cases} c_i(\mathbf{x})_- &= c_i(\mathbf{x}), \qquad\qquad i \in \mathcal{E} \\ c_i(\mathbf{x})_- &= |\min\{0, c_i(\mathbf{x})\}|, \quad i \in \mathcal{I} \end{cases} \tag{3.21}$$

Consider simple penalty function

$$P_\sigma(\mathbf{x}) = f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{3.22}$$

Denote $\mathbf{x}(\sigma)$ as the solution to unconstrained problem $\min P_\sigma(\mathbf{x})$, we have the following lemma:

Lemma 3.3.1 (Penalty method)**.** *If* $\mathbf{x}(\sigma)$ *is a feasible point of nonlinear constrained problem* (3.17)*, then* $\mathbf{x}(\sigma)$ *aslo is the solution to* (3.17)*.*

Proof. From the definition of penalty function, we have $P(\mathbf{x}) = f(\mathbf{x}), \mathbf{x} \in \mathcal{S}$. If $\mathbf{x}(\sigma)$ is the solution to $\min P(\mathbf{x})$, i.e.,

$$P(\mathbf{x}(\sigma)) \leq P(\mathbf{x}_0),\ \forall \mathbf{x}_0 \in \mathbb{R}^n \tag{3.23}$$

$$f(\mathbf{x}(\sigma)) \leq f(\mathbf{x}_0),\ \forall \mathbf{x}_0 \in \mathcal{S} \tag{3.24}$$

that is, $\mathbf{x}(\sigma)$ is the solution to (3.17).

□

Algorithm 6: Penalty Method Algorithm

Data: Cost function f
$x^{(0)} \in \mathbb{R}^n$, $\sigma_0 > 0$, $\beta > 1$, $\epsilon > 0$, $k := 0$;
while $\parallel \mathbf{c}(\mathbf{x}(\sigma_{k-1}))_- \parallel \geq \epsilon$ **do**
  solve the subproblem $\min_{\mathbf{x}\in\mathbb{R}^n} P_{\sigma_k}(\mathbf{x})$ to get the solution $\mathbf{x}(\sigma_k)$;
  $\mathbf{x}^{(k+1)} = \mathbf{x}(\sigma_k)$, $\sigma_{k+1} = \beta\sigma_k$;
  $k := k+1$;
end
return: $\mathbf{x}(\sigma_{k-1})$

Theorem 3.3.2 (Convergence of Penalty method)**.** *If* $\epsilon > \min_{\mathbf{x}\in\mathbb{R}^n} \parallel \mathbf{c}(\mathbf{x})_- \parallel$*, then the algorithm can terminate in finite steps.*

Lemma 3.3.3. *Let* $\sigma_{k+1} > \sigma_k > 0$*, then we have* $P_{\sigma_k}(\mathbf{x}(\sigma_k)) \leq P_{\sigma_{k+1}}(\mathbf{x}(\sigma_{k+1}))$*,* $\| \mathbf{c}(\mathbf{x}(\sigma_k))_- \| \geq \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|$*,* $f(\mathbf{x}(\sigma_k)) \leq f(\mathbf{x}(\sigma_{k+1}))$*.*

Proof.

$$
\begin{aligned}
P_{\sigma_{k+1}}(\mathbf{x}(\sigma_{k+1})) &= f(\mathbf{x}(\sigma_{k+1})) + \sigma_{k+1} \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 & (3.25)\\
&\geq f(\mathbf{x}(\sigma_{k+1})) + \sigma_k \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 & (3.26)\\
&\geq \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \sigma_k \| \mathbf{c}(\mathbf{x})_- \|^2 & (3.27)\\
&= P_{\sigma_k}(\mathbf{x}(\sigma_k)) & (3.28)
\end{aligned}
$$

From the definition, we have

$$
\begin{aligned}
& f(\mathbf{x}(\sigma_k)) + \sigma_{k+1} \| \mathbf{c}(\mathbf{x}(\sigma_k))_- \|^2 & (3.29)\\
\geq & f(\mathbf{x}(\sigma_{k+1})) + \sigma_{k+1} \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 & (3.30)\\
\geq & f(\mathbf{x}(\sigma_{k+1})) + \sigma_k \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 & (3.31)\\
\geq & f(\mathbf{x}(\sigma_k)) + \sigma_k \| \mathbf{c}(\mathbf{x}(\sigma_k))_- \|^2 & (3.32)
\end{aligned}
$$

From the inequalities above, we have

$$
\begin{aligned}
& \sigma_k(\| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 - \| \mathbf{c}(\mathbf{x}(\sigma_k))_- \|^2) & (3.33)\\
\leq & f(\mathbf{x}(\sigma_{k+1})) - f(\mathbf{x}(\sigma_k)) & (3.34)\\
\leq & \sigma_{k+1}(\| \mathbf{c}(\mathbf{x}(\sigma_k))_- \|^2 - \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2) & (3.35)
\end{aligned}
$$

So that

$$\| \mathbf{c}(\mathbf{x}(\sigma_k))_- \| \geq \| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \| \tag{3.36}$$

Then

$$0 \leq \sigma_k(\| \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \|^2 - \| \mathbf{c}(\mathbf{x}(\sigma_k))_- \|^2) \leq f(\mathbf{x}(\sigma_{k+1})) - f(\mathbf{x}(\sigma_k)) \tag{3.37}$$

i.e.,

$$f(\mathbf{x}(\sigma_{k+1})) \geq f(\mathbf{x}(\sigma_k)) \tag{3.38}$$

$\square$

Lemma 3.3.4. *Denote* $\bar{\mathbf{x}}$ *as the solution to problem* (3.17)*, then for all* $\sigma_k > 0$*,*

$$f(\bar{\mathbf{x}}) \geq P_{\sigma_k}(\mathbf{x}(\sigma_k)) \geq f(\mathbf{x}(\sigma_k)) \tag{3.39}$$

Proof. For all $\sigma_k > 0$,

$$f(\bar{\mathbf{x}}) = \min_{\mathbf{x}\in\mathbb{R}^n} \lim_{\sigma\to\infty} f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{3.40}$$

$$\geq \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \sigma_k \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{3.41}$$

$$= f(\mathbf{x}(\sigma_k)) + \sigma_k \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2 \tag{3.42}$$

$$\geq f(\mathbf{x}(\sigma_k)) \tag{3.43}$$

□

Lemma 3.3.5. *Let* $\delta = \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel$*, then* $\mathbf{x}(\sigma)$ *is also the solution to the problem*

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & \parallel \mathbf{c}(\mathbf{x})_- \parallel \leq \delta \end{array} \tag{3.44}$$

Proof. The problem is equivalent to

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & \parallel \mathbf{c}(\mathbf{x})_- \parallel \leq \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel \end{array} \tag{3.45}$$

$$f(\mathbf{x}(\sigma)) + \sigma \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel^2 = \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{3.46}$$

Then for all $\mathbf{x} \in \mathbb{R}^n$, we have

$$f(\mathbf{x}(\sigma)) + \sigma \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel^2 \leq f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{3.47}$$

$$f(\mathbf{x}(\sigma)) - f(\mathbf{x}) \leq \sigma(\parallel \mathbf{c}(\mathbf{x})_- \parallel^2 - \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel^2) \tag{3.48}$$

That is, if $\parallel \mathbf{c}(\mathbf{x})_- \parallel \leq \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel$, then

$$f(\mathbf{x}(\sigma)) - f(\mathbf{x}) \leq \sigma(\parallel \mathbf{c}(\mathbf{x})_- \parallel^2 - \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel^2) \leq 0 \tag{3.49}$$

i.e., for all $\mathbf{x} \in \mathbb{R}^n$, $f(\mathbf{x}(\sigma)) \leq f(\mathbf{x})$.

□

3.3.3 Argumented Lagrange function method

3.3.3.1 Revisit Penalty method

Consider equality constrained problem

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & \mathbf{c}(\mathbf{x}) = 0 \end{array} \tag{3.50}$$

The Lagrange function of (3.50) is

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T \mathbf{c}(\mathbf{x}) \tag{3.51}$$

From K-T condition, we have for global optimal point $\mathbf{x}^*$,

$$\begin{cases} \bigtriangledown_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \lambda^*) = 0 \\ \bigtriangledown_{\lambda}\mathcal{L}(\mathbf{x}^*, \lambda^*) = 0 \end{cases} \tag{3.52}$$

i.e., $\mathbf{x}^*$ is a stable point of $\mathcal{L}(\mathbf{x}, \lambda)$. Notice that

$$\bigtriangledown \mathbf{x}\mathcal{L}(\mathbf{x}^*, \lambda^*) = \bigtriangledown f(\mathbf{x}^*) - \sum_i \lambda_i^* \bigtriangledown c_i(\mathbf{x}^*) \tag{3.53}$$

For the corresponding penalty function

$$P_\sigma(\mathbf{x}) = f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x}) \parallel^2 \tag{3.54}$$

we have the K-T condition is

$$\bigtriangledown P_\sigma(\mathbf{x}^*) = \bigtriangledown f(\mathbf{x}^*) + 2\sigma \mathbf{c}^T(\mathbf{x}^*) \bigtriangledown \mathbf{c}(\mathbf{x}^*) \tag{3.55}$$

$$= \bigtriangledown f(\mathbf{x}^*) + \sum_i 2\sigma c_i(\mathbf{x}^*) \bigtriangledown c_i(\mathbf{x}^*) = 0 \tag{3.56}$$

If we want (3.56) to be a good approximation of (3.53), i.e.,

$$\lambda_i^* \approx -2\sigma c_i(\mathbf{x}^*) \tag{3.57}$$

Notice that $c_i(\mathbf{x}^*) \approx 0$, so we need $|\sigma| \to \infty$.

3.3.3.2 Argumented Lagrange function method

Consider *Argumented Lagrange function*

$$\min_{\mathbf{x}, \lambda} P(\mathbf{x}, \lambda, \sigma) = \mathcal{L}(\mathbf{x}, \lambda) + \frac{\sigma}{2} \parallel \mathbf{c}(\mathbf{x}) \parallel^2 \tag{3.58}$$

The K-T condition of the function is

$$\begin{cases} \bigtriangledown_{\mathbf{x}} P(\mathbf{x}^*, \lambda^*, \sigma) = 0 \\ \bigtriangledown_{\lambda} P(\mathbf{x}^*, \lambda^*, \sigma) = 0 \end{cases} \tag{3.59}$$

$$\bigtriangledown_{\lambda} P(\mathbf{x}^*, \lambda^*, \sigma) = \mathbf{c}(\mathbf{x}) = 0 \tag{3.60}$$

$$\bigtriangledown_{\mathbf{x}} P(\mathbf{x}^*, \lambda^*, \sigma) = \bigtriangledown f(\mathbf{x}^*) - \sum_i (\lambda_i^* - \sigma c_i(\mathbf{x}^*)) \bigtriangledown c_i(\mathbf{x}^*) \tag{3.61}$$

$$= \bigtriangledown f(\mathbf{x}^*) - \sum_i \lambda_i^* \bigtriangledown c_i(\mathbf{x}^*) = 0 \tag{3.62}$$

i.e., the K-T condition of P is similar to the original problem (3.50).

Theorem 3.3.6. *Suppose $\mathbf{x}^*$ and λ^* satisfy the K-T condition of (3.50), then there exists $\bar{\sigma}$ such that when $\sigma > \bar{\sigma}$, $\mathbf{x}^*$ is the strict local minima of $P(\mathbf{x}, \lambda^*, \sigma)$.*

Proof. Appearently if $\mathbf{x}^*$ and λ^* satisfy the K-T condition of (3.50), then $\mathbf{x}^*$ and λ^* also satisfy the K-T condition of (3.58).

For (3.58), we can always find $\bar{\sigma}$ when $\sigma > \bar{\sigma}$, the problem is convex. In this case, the K-T condition is sufficient and necessary condition of optimal points. □

However, the optimal value λ^* remains unknown.

Algorithm 7: Argumented Lagrange Algorithm

Data: Cost function f

$x^{(0)} \in \mathbb{R}^n$, $\sigma_0 > 0$, $\alpha > 1$, $0 < \beta < 1$, $\epsilon > 0$, $k := 0$;

while $\| \mathbf{c}(\mathbf{x}^{(k)} \| \geq \epsilon$ **do**

  $\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}\in\mathbb{R}^n} P(\mathbf{x}, \lambda^{(k)}, \sigma)$;

  $\lambda^{(k+1)} = \lambda^{(k)} - \sigma \mathbf{c}(\mathbf{x}^{(k+1)})$;

  if $\| \mathbf{c}(\mathbf{x}^{(k+1)} \| / \| \mathbf{c}(\mathbf{x}^{(k)} \| \geq \beta$ **then**

    $\sigma := \alpha\sigma$

  end

  $k := k + 1$;

end

return: $\mathbf{x}^{(k)}$

3.3.4 Barrier method

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & g_i(\mathbf{x}) \geq 0, i = 1, ..., m \end{array} \tag{3.63}$$

We use $\mathbf{int}S$ to denote the interior of feasible region, where $S = \{\mathbf{x} \mid g_i(\mathbf{x}) \geq 0, i = 1, ..., m\}$. Define *Barrier function*

$$B(\mathbf{x}, \theta) = f(\mathbf{x}) + \theta\psi(\mathbf{x}) \tag{3.64}$$

Where barrier factor θ is a small positive number, $\psi(\mathbf{x})$ is a continuous function. When $\mathbf{x} \to \partial S$, $\psi(\mathbf{x}) \to +\infty$. We can derive the approximate solution to the original problem (3.63)

$$\begin{array}{ll} \min & B(\mathbf{x}, \theta) \\ s.t. & \mathbf{x} \in \mathbf{int}S \end{array} \tag{3.65}$$

Algorithm 8: Barrier Algorithm

Data: Cost function f, feasible region S

$x^{(0)} \in \mathbf{int}S$, $\theta_0 > 0$, $0 < \beta < 1$, $\epsilon > 0$, $k := 0$;

while $\theta_k \psi(\mathbf{x}^{(k)} \geq \epsilon$ **do**

  $\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x} \in \mathbf{int}S} f(\mathbf{x}) + \theta_k \psi(\mathbf{x})$;

  $\theta_{k+1} := \beta\theta_k$;

  $k := k + 1$;

end

return: $\mathbf{x}^{(k)}$

Theorem 3.3.7. *Suppose* $\theta_k > \theta_{k+1} > 0$*, denote* $\mathbf{x}(\theta) = \arg\min_{\mathbf{x}} B(\mathbf{x}, \theta)$*, then*

$$B(\mathbf{x}(\theta_k), \theta_k) \geq B(\mathbf{x}(\theta_{k+1}), \theta_{k+1}) \tag{3.66}$$

$$\psi(\mathbf{x}(\theta_k)) \leq \psi(\mathbf{x}(\theta_{k+1})) \tag{3.67}$$

$$f(\mathbf{x}(\theta_k)) \geq f(\mathbf{x}(\theta_{k+1})) \tag{3.68}$$

Proof. Similar to Proof of Lemma (3.3.3),

$$\begin{aligned}
B(\mathbf{x}(\theta_k), \theta_k) &= f(\mathbf{x}(\theta_k)) + \theta_k \psi(\mathbf{x}(\theta_k)) &(3.69)\\
&\geq f(\mathbf{x}(\theta_k)) + \theta_{k+1} \psi(\mathbf{x}(\theta_k)) &(3.70)\\
&\geq \min_{\mathbf{x} \in \mathbf{int}S} f(\mathbf{x}) + \theta_{k+1}\psi(\mathbf{x}) &(3.71)\\
&= B(\mathbf{x}(\theta_{k+1}), \theta_{k+1}) &(3.72)
\end{aligned}$$

From

$$\begin{aligned}
&f(\mathbf{x}(\theta_{k+1})) + \theta_k \psi(\mathbf{x}(\theta_{k+1})) &(3.73)\\
\geq &f(\mathbf{x}(\theta_k)) + \theta_k \psi(\mathbf{x}(\theta_k)) &(3.74)\\
\geq &f(\mathbf{x}(\theta_k)) + \theta_{k+1} \psi(\mathbf{x}(\theta_k)) &(3.75)\\
\geq &f(\mathbf{x}(\theta_{k+1})) + \theta_{k+1} \psi(\mathbf{x}(\theta_{k+1})) &(3.76)
\end{aligned}$$

we have

$$\theta_k(\psi(\mathbf{x}(\theta_k)) - \psi(\mathbf{x}(\theta_{k+1}))) \leq f(\mathbf{x}(\theta_{k+1})) - f(\mathbf{x}(\theta_k)) \leq \theta_{k+1}(\psi(\mathbf{x}(\theta_k)) - \psi(\mathbf{x}(\theta_{k+1}))) \tag{3.77}$$

notice that $\theta_k > \theta_{k+1} > 0$, so

$$\psi(\mathbf{x}(\theta_k)) \leq \psi(\mathbf{x}(\theta_{k+1})) \tag{3.78}$$

$$f(\mathbf{x}(\theta_{k+1})) - f(\mathbf{x}(\theta_k)) \leq \theta_{k+1}(\psi(\mathbf{x}(\theta_k)) - \psi(\mathbf{x}(\theta_{k+1}))) \leq 0 \tag{3.79}$$

$$f(\mathbf{x}(\theta_{k+1})) \leq f(\mathbf{x}(\theta_k)) \tag{3.80}$$

□

4 Convex Optimization

4.1 Convex set

4.1.1 Affine set

Definition 4.1.1 (Affine set)**.** *A set $\mathcal{C} \subset \mathbb{R}^n$ is affine if $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ and $\theta \in \mathbb{R}$, we have*

$$\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{C} \tag{4.1}$$

Definition 4.1.2 (Affine hull)**.** *The set of all affine combinations of points in some set $\mathcal{C} \subset \mathbb{R}^n$ is called the* affine hull *of $\mathcal{C}$, denoted $\mathbf{aff}\mathcal{C}$:*

$$\mathbf{aff}\mathcal{C} = \{\sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_1, ..., \mathbf{x}_k \in \mathcal{C}, \theta_1 + ... + \theta_k = 1\} \tag{4.2}$$

Remark 2. *The affine hull is the smallest affine set that contains $\mathcal{C}$.*

Proof. For any affine set $\mathcal{A}$ contains $\mathcal{C}$, we have

$$\sum_{i=1}^{k} \theta_i \mathbf{x}_i \in \mathcal{A}, \forall \mathbf{x}_1, ..., \mathbf{x}_k \in \mathcal{C}, \theta_1 + ... + \theta_k = 1 \tag{4.3}$$

i.e., $\mathbf{aff}\mathcal{C} \subset \mathcal{A}$. □

4.1.2 Convex set

Definition 4.1.3 (Convex set)**.** *A set $\mathcal{C} \subset \mathbb{R}^n$ is convex if $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ and $0 \le \theta \le 1$, we have*

$$\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{C} \tag{4.4}$$

Definition 4.1.4 (Convex hull)**.** *The set of all convex combinations of points in some set $\mathcal{C} \subset \mathbb{R}^n$ is called the* convex hull *of $\mathcal{C}$, denoted $\mathbf{conv}\mathcal{C}$:*

$$\mathbf{conv}\mathcal{C} = \{\sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_1, ..., \mathbf{x}_k \in \mathcal{C}, \theta_i \ge 0, \theta_1 + ... + \theta_k = 1\} \tag{4.5}$$

Remark 3. *The convex hull is the smallest convex set that contains $\mathcal{C}$.*

4.1.3 Cone

Definition 4.1.5 (Cone)**.** *A set $\mathcal{C}$ is called a* cone*, if $\forall \mathbf{x} \in \mathcal{C}$ and $\theta \ge 0$ we have $\theta \mathbf{x} \in \mathcal{C}$. A set $\mathcal{C}$ is called a* convex cone *if it is convex and a cone, i.e., $\forall \mathbf{x}_1, \mathbf{x})_2 \in \mathcal{C}$ and $\theta_1, \theta_2 \ge 0$, we* *have*

$$\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in \mathcal{C} \tag{4.6}$$

Definition 4.1.6 (Conic hull)**.** *The conic hull of set $\mathcal{C}$ is the set of all conic combinations of points in $\mathcal{C}$, i.e.,*

$$\{\sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in \mathcal{C}, \theta_i \geq 0, i = 1, ..., k\} \tag{4.7}$$

4.1.4 Proper cones and generalized inequalities

4.2 Convex function

Definition 4.2.1 (Convex function)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* convex *if $\mathbf{dom} f$ is a convex set and if $\forall x, y \in \mathbf{dom} f$ and θ with $0 \leq \theta \leq 1$, we have*

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2) \tag{4.8}$$

4.2.1 First order condition

Suppose f is differentiable

Theorem 4.2.1. *Function f is convex if and only if $\mathbf{dom} f$ is a convex set and for $\forall x, y \in \mathbf{dom} f$, the following holds:*

$$f(y) \geq f(x) + \bigtriangledown f(x)^T (y - x) \tag{4.9}$$

Remark 4. *If $\bigtriangledown f(x^*) = 0$, then for $\forall y \in \mathbf{dom} f$, $f(y) \geq f(x^*)$, i.e., x^* is the global minimizer of f.*

4.2.2 Second order condition

Suppose f is twice differentiable

Theorem 4.2.2. *Function f is convex if and only if $\mathbf{dom} f$ is a convex set and for $\forall x \mathbf{dom} f$, the following holds:*

$$\bigtriangledown^2 f(x) \succeq 0 \tag{4.10}$$

Remark 5. *If $\bigtriangledown^2 f(x) \succ 0$ for $\forall x \mathbf{dom} f$, then f is* strictly convex.

4.2.3 Properties of Convex functions

4.2.3.1 Jensen's Inequality

Theorem 4.2.3 (Jensen's Inequality)**.** *If f is convex, $x_1, ..., x_k \in \mathbf{dom} f$, and $\theta_1, ..., \theta_k \geq 0$ with $\theta_1 + ... + \theta_k = 1$, then*

$$f(\theta_1 x_1 + ... + \theta_k x_k) \leq \theta_1 f(x_1) + ... + \theta_k f(x_k) \tag{4.11}$$

4.2.3.2 Operations that preserve convexity

Nonnegative weighted sums If $f_1, ..., f_m$ are covex and $w_1, ..., w_m \geq 0$, then

$$f = w_1 f_1 + ... + w_m f_m \tag{4.12}$$

is convex.

If $f(x, y)$ is convex w.r.t x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y) dy \tag{4.13}$$

is convex w.r.t x.

Composition with an affine mapping Suppose $f : \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{n \times m}$, and $\mathbf{b} \in \mathbb{R}$. Define $g : \mathbb{R}^m \to \mathbb{R}$ by

$$g(\mathbf{x}) = f(A\mathbf{x} + \mathbf{b}) \tag{4.14}$$

with $\mathbf{dom} g = \{\mathbf{x} \mid A\mathbf{x} + \mathbf{b} \in \mathbf{dom} f\}$. Then if f is convex, so is g.

Pointwise maximum If f_1 and f_2 are convex functions, then

$$f(x) = \max\{f_1(x), f_2(x)\} \tag{4.15}$$

with $\mathbf{dom} f = \mathbf{dom} f_1 \cap \mathbf{dom} f_2$ is also convex.

If $f(x, y)$ is convex w.r.t x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y) \tag{4.16}$$

is convex in x, where

$$\mathbf{dom} g = \{x \mid (x, y) \in \mathbf{dom} f, \forall y \in \mathcal{A}, \sup f(x, y) < \infty\} \tag{4.17}$$

4.2.4 Quasi-convex function

Definition 4.2.2 (Quasi-convex function)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ sucha that its domain and all its sublevel sets*

$$S_\alpha = \{x \in \mathbf{dom} f \mid f(x) \leq \alpha\}, \alpha \in \mathbb{R} \tag{4.18}$$

are convex, then f is quasi-convex.

4.3 Convex optimization

A *convex optimization problem* is one of the form

$$\begin{array}{ll} \min & f_0(\mathbf{x}) \\ s.t. & f_i(\mathbf{x}) \leq 0, \quad i = 1, ..., m \\ & a_j^T \mathbf{x} = b_j, \quad j = 1, ..., p \end{array} \tag{4.19}$$

where $f_0, ..., f_m$ are convex functions.

Remark 6. *The equality constraint is linear if the problem is convex.*

Proof. For equality constraint

$$\mathbf{c}(\mathbf{x}) = 0 \tag{4.20}$$

we can rewrite it into

$$\mathbf{c}(\mathbf{x}) \leq 0 \tag{4.21}$$

$$-\mathbf{c}(\mathbf{x}) \leq 0 \tag{4.22}$$

Due to the convexity of the problem, both $\mathbf{c}(\mathbf{x})$ and $-\mathbf{c}(\mathbf{x})$ are convex. i.e., $\mathbf{c}(\mathbf{x})$ is linear. □

4.3.1 Optimal condition

Theorem 4.3.1 (Optimal condition)**.** *Suppose (4.19) is differentiable. Let S denote the feasible set, then $\mathbf{x}^*$ is optimal if and only if $\mathbf{x}^* \in S$ and*

$$\triangledown f_0(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \geq 0, \forall y \in S \tag{4.23}$$

Proof. If $\mathbf{x}^*$ is optimal, then we can easily derive (4.23).

If (4.23) stands, then from Theorem 4.2.1,

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \triangledown f_0(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \geq 0, \forall y \in S \tag{4.24}$$

□

Lemma 4.3.2. *For convex problem with equality constraints only, i.e.,*

$$\begin{array}{ll} \min & f_0(\mathbf{x}) \\ s.t. & A(\mathbf{x}) = \mathbf{b} \end{array} \tag{4.25}$$

the optimal condition can be expressed as

$$\triangledown f_0(\mathbf{x})^T \mathbf{u} \geq 0, \forall \mathbf{u} \in \mathcal{N}(A) \tag{4.26}$$

in other words,

$$\bigtriangledown f_0(\mathbf{x}) \perp \mathcal{N}(A) \tag{4.27}$$

Proof. From Theorem 4.3.1, we have $\mathbf{x}^*$ is optimal if and only if $A\mathbf{x} = \mathbf{b}$, for $\forall \mathbf{y}$ such that $A\mathbf{y} = \mathbf{b}$,

$$\bigtriangledown f_0(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \geq 0 \tag{4.28}$$

i.e., $A(\mathbf{y} - \mathbf{x}) = 0$. Let $\mathbf{u} = \mathbf{y} - \mathbf{x}$, then

$$\bigtriangledown f_0(\mathbf{x})^T \mathbf{u} \geq 0, \forall \mathbf{u} \in \mathcal{N}(A) \tag{4.29}$$

further, if $\mathbf{u} \in \mathcal{N}(A)$, then, $-\mathbf{u} \in \mathcal{N}(A)$, so we have

$$\bigtriangledown f_0(\mathbf{x})^T \mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{N}(A) \tag{4.30}$$

i.e.,

$$\bigtriangledown f_0(\mathbf{x}) \perp \mathcal{N}(A) \tag{4.31}$$

$\square$

Lemma 4.3.3 (Global optimality)**.** *Any locally optimal point is also globally optimal in convex optimization problems.*

4.3.2 Common convex optimizations

4.3.2.1 Linear optimization

A general *linear program* (LP) has the form

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} + d \\ s.t. \quad & G\mathbf{x} \leq \mathbf{h} \\ & A\mathbf{x} = \mathbf{b} \end{aligned} \tag{4.32}$$

where $G \in \mathbb{R}^{m\times n}$ and $A \in \mathbb{R}^{p\times n}$.

4.3.2.2 Quadratic optimization

A general *quadratic program* (QP) has the form

$$\begin{aligned} \min \quad & \tfrac{1}{2}\mathbf{x}^T P\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ s.t. \quad & G\mathbf{x} \leq \mathbf{h} \\ & A\mathbf{x} = \mathbf{b} \end{aligned} \tag{4.33}$$

where $P \in \mathbf{S}_+^n$, $G \in \mathbb{R}^{m\times n}$ and $A \in \mathbb{R}^{p\times n}$.

Quadratically constrained quadratic program

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{x}^T P_0 \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \\ s.t. & \frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 1, ..., m \\ & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.34}$$

where $P_i \in \mathbf{S}_+^n, i = 0, ..., m$, the problem is called a *quadratically constrained quadratic program* (QCQP).

Second-order cone program

$$\begin{array}{ll} \min & \mathbf{f}^T \mathbf{x} \\ s.t. & \| A_i \mathbf{x} + \mathbf{b}_i \| \leq \mathbf{c}_i^T \mathbf{x} + \mathbf{d}_i, \quad i = 1, ..., m \\ & F\mathbf{x} = \mathbf{g} \end{array} \tag{4.35}$$

Lemma 4.3.4. *Any QCQP problem can be formulated as a SOCP problem.*

Proof. The QCQP problem is equivalent to

$$\begin{array}{ll} \min & -r_0 \\ s.t. & \frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 0, ..., m \\ & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.36}$$

Then we need to prove that (4.36) can be formulated as (4.35).

$$\frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0 \tag{4.37}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + 2(\mathbf{q}_i^T \mathbf{x} + r_i) \leq 0 \tag{4.38}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + 2(\mathbf{q}_i^T \mathbf{x} + r_i) + (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{4.39}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + (\mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{4.40}$$

Since P_i is positive semi-definite, $P_i = A_i^T A_i$, then

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + (\mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{4.41}$$

$$\Leftrightarrow \| A_i \mathbf{x} \|^2 + \| \mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2} \|^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{4.42}$$

Let

$$A_i' = \begin{pmatrix} A \\ \mathbf{q}^T \end{pmatrix} \tag{4.43}$$

$$\mathbf{b}_i = \begin{pmatrix} \mathbf{0}_{n \times 1} \\ r_i + \frac{1}{2} \end{pmatrix} \tag{4.44}$$

From (4.37) and $\mathbf{x}^T P_i \mathbf{x} \geq 0$, we can derive that $\mathbf{q}_i^T \mathbf{x} + r_i \leq 0$, then, $\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2} \leq 0$.

Then (4.42) can be formulated as

$$\| A_i'\mathbf{x} + \mathbf{b}_i \|^2 \leq (\mathbf{q}_i^T\mathbf{x} + r_i - \frac{1}{2})^2 \tag{4.45}$$

$$\Leftrightarrow \| A_i'\mathbf{x} + \mathbf{b}_i \| \leq -(\mathbf{q}_i^T\mathbf{x} + r_i - \frac{1}{2}) \tag{4.46}$$

□

4.3.3 Lagrange dual problem

Consider optimization problem

$$\begin{array}{lll} \min & f_0(\mathbf{x}) & \\ s.t. & f_i(\mathbf{x}) \leq 0, & i = 1, ..., m \\ & h_j(\mathbf{x}) = 0, & j = 1, ..., p \end{array} \tag{4.47}$$

, Denote the optimal value of Problem 4.47 by v^*, but we *do not* assume the problem is convex.

Recall the K-T conditions in Chapter 1, we can define *Lagrangian* by

Definition 4.3.1 (Lagrangian)**.** *The Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ *associated with the Problem 4.47 is*

$$L(\mathbf{x}, \lambda, \nu) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j h_i(\mathbf{x}) \tag{4.48}$$

with $\mathbf{dom}L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$.

Refer to λ_i as the *Lagrange multiplier* associated with the ith inequality constraint $f_i(\mathbf{x}) \leq 0$.

Refer to ν_j as the Lagrange multiplier associated with the jth inequality constraint $h_j(\mathbf{x}) = 0$.

The vectors λ and ν are called the *dual variables* or *Lagrange multiplier vectors*.

Definition 4.3.2. *The* Lagrange dual function *of Problem 4.47 is*

$$g(\lambda, \nu) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \lambda, \nu) = \inf_{\mathbf{x} \in \mathcal{D}} \Big(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j h_i(\mathbf{x}) \Big) \tag{4.49}$$

Notice that for all $\mathbf{x} \in \mathcal{D}$, $L(\mathbf{x}, \lambda, \nu)$ is affine w.r.t λ, ν, that is, concave w.r.t λ, ν. Recall that *Pointwise maximum* operation can preserve convexity, i.e., *Pointwise infimum* can preserve concavity. So the Lagrange dual function is concave.

4.3.3.1 Lower bounds optimal value

Theorem 4.3.5. *For any* $\lambda \geq 0$ *and any* ν*, we have*

$$g(\lambda, \nu) \leq v^* \tag{4.50}$$

Proof. Denote the optimal point of Problem 4.47 as $\mathbf{x}^*$, then appearently $\mathbf{x}^*$ is a feasible point, i.e.,

$$\begin{cases} f_i(\mathbf{x}) \leq 0, & i = 1, ..., m \\ h_j(\mathbf{x}) = 0, & j = 1, ..., p \end{cases} \tag{4.51}$$

then we have

$$g(\lambda, \nu) \leq L(\mathbf{x}^*, \lambda, \nu) = f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^*) + \sum_{j=1}^{p} \nu_j h_i(\mathbf{x}^*) \leq f_0(\mathbf{x}^*) = v^* \tag{4.52}$$

□

We refer to a pair (λ, ν) with $\lambda \geq 0$ and $(\lambda, \nu) \in \mathbf{dom} g$ as *dual feasible*.

Linear approximation interpretation Define functions

$$I_{\text{-}}(u) = \begin{cases} 0, & u \leq 0 \\ +\infty, & u > 0 \end{cases} \tag{4.53}$$

$$I_0(u) = \begin{cases} 0, & u = 0 \\ +\infty, & u \neq 0 \end{cases} \tag{4.54}$$

Then the Problem 4.47 is equivalent to

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) + \sum_{i=1}^{m} I_{\text{-}}(f_i(\mathbf{x})) + \sum_{j=1}^{p} I_0(h_j(\mathbf{x})) \tag{4.55}$$

Appearently (4.49) is a softer version of (4.55), so Theorem 4.3.5 holds.

4.3.3.2 The Lagrange dual problem

To attain the best lower bound of v^*, we can solve the following optimization problem

$$\begin{array}{ll} \max & g(\lambda, \nu) \\ s.t. & \lambda \geq 0 \end{array} \tag{4.56}$$

This problem is called *Lagrange dual problem* associated with Problem 4.47. Correspondingly, Problem 4.47 is called the *primal problem*.

We refer (λ^*, ν^*) as *dual optimal* or *optimal Lagrange multipliers* if they are optimal for Problem 4.56.

Notice that the Lagrange dual problem is convex whether the primal problem is convex or not.

4.3.3.3 Weak duality

For the optimal value of Lagrange dual problem 4.56 g^*, we have

$$g^* \le v^* \tag{4.57}$$

This property is called *weak duality*.

$v^* - g^*$ is the *optimal duality gap* of the primal problem.

4.3.3.4 Strong duality

For the optimal value of Lagrange dual problem 4.56 g^*, if

$$g^* = v^* \tag{4.58}$$

holds, then we say that *weak duality* holds.

Definition 4.3.3 (Strictly feasible)**.** *For a feasible point* $\mathbf{x}$*, if*

$$f_i(\mathbf{x}) < 0, i = 1, ..., m \tag{4.59}$$

$$A\mathbf{x} = \mathbf{b} \tag{4.60}$$

holds, then we called $\mathbf{x}$ *is* strictly feasible*.*

Definition 4.3.4 (Relative interior)**.** *The* relative interior *of set* $\mathcal{D}$ *is*

$$\textbf{relint}\mathcal{D} = \{\mathbf{x} \in \mathcal{D} \mid \exists r > 0, B(\mathbf{x}, r) \cap \textbf{aff}\mathcal{D} \subset \mathcal{D}\} \tag{4.61}$$

Theorem 4.3.6 (Slater's condition)**.** *If there exists an* $\mathbf{x} \in \textbf{relint}\mathcal{D}$ *that is strictly feasible, then strong duality holds (and the problem is convex).*

4.3.3.5 Complementary slackness

If strong duality holds, i.e.,

$$\begin{aligned} f_0(\mathbf{x}^*) &= g(\lambda^*, \nu^*) && (4.62)\\ &= \inf_{\mathbf{x}} \Big(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j^* h_i(\mathbf{x}) \Big) && (4.63)\\ &\le f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{j=1}^{p} \nu_j^* h_i(\mathbf{x}^*) && (4.64)\\ &\le f_0(\mathbf{x}^*) && (4.65) \end{aligned}$$

Notice that $\lambda_i^* \ge 0$ and $f_i(\mathbf{x}^*) \le 0$, we have

$$\lambda_i^* f_i(\mathbf{x}^*) = 0, i = 1, ..., m \tag{4.66}$$

This condition is known as *complementary slackness*.

4.3.4 KKT optimality conditions

4.3.4.1 KKT optimality conditions for nonconvex problems

Consider optimization problem

$$\begin{array}{lll} \min & f_0(\mathbf{x}) & \\ s.t. & f_i(\mathbf{x}) \leq 0, & i = 1, ..., m \\ & h_j(\mathbf{x}) = 0, & j = 1, ..., p \end{array} \tag{4.67}$$

Assume that $f_0, ..., f_m, h_1, ..., h_p$ are differentiable. Let $\mathbf{x}^*$ and (λ^*, ν^*) be any primal and dual optimal points with *zero* duality gap.

Summarize the optimal conditions, we have

$$(KKT) \begin{cases} f_i(\mathbf{x}^*) \leq 0, i = 1, ..., m \\ h_j(\mathbf{x}^*) = 0, j = 1, ..., p \\ \lambda_i^* \geq 0, i = 1, ..., m \\ \lambda_i^* f_i(\mathbf{x}^*) = 0, i = 1, ..., m \\ \bigtriangledown f_0(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \bigtriangledown f_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \bigtriangledown h_i(\mathbf{x}^*) = 0 \end{cases} \tag{4.68}$$

Recall the K-T conditions (1.34) in Chapter 1, we can see that the assumption is a little bit different.

The relation between K-T condition and Slater's condition?

4.3.4.2 KKT optimality conditions for convex problems

If Problem 4.67 is convex, then the KKT conditions are also sufficient for primal and dual optimality. That is to say, if f_i are convex and h_i are affine, then any points satisfy the KKT conditions are primal and dual potimal points *with zero duality gap*.

4.4 Newton method for equality constrained problems

4.4.1 Problem formulation

A convex optimization problem with equality constraints

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.69}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and twice continuously differentiable, and $A \in \mathbb{R}^{p \times n}$ with $\mathbf{rank} A = p < n$. We assume that an optimal point $\mathbf{x}^*$ exists and $v^* = f(\mathbf{x}^*)$.

Recall the Newton Method for unconstrained problems, i.e., find the minima of the quadratic approximation model.

$$\begin{array}{ll} \min_{\mathbf{s}} & f(\mathbf{x} + \mathbf{s}) \\ s.t. & A(\mathbf{x} + \mathbf{s}) = \mathbf{b} \end{array} \tag{4.70}$$

From K-T condition, we have

$$\begin{cases} \bigtriangledown f(\mathbf{x}+\mathbf{s}) + A^T\lambda = 0 \\ A(\mathbf{x}+\mathbf{s}) = \mathbf{b} \end{cases} \tag{4.71}$$

Similarly we derive the Newton step in this case.

4.4.2 Newton method with feasible start

This method requires a feasible initial point. The Newton step is the solution of the problem

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{s}^T \bigtriangledown^2 f(\mathbf{x})\mathbf{s} + \bigtriangledown f(\mathbf{x})^T\mathbf{s} + f(\mathbf{x}) \\ s.t. & A(\mathbf{x}+\mathbf{s}) = \mathbf{b} \end{array} \tag{4.72}$$

Notice that the initial point $\mathbf{x} \in \mathcal{S}$, so that $A\mathbf{x} = \mathbf{b}$, then we have

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{s}^T \bigtriangledown^2 f(\mathbf{x})\mathbf{s} + \bigtriangledown f(\mathbf{x})^T\mathbf{s} + f(\mathbf{x}) \\ s.t. & A\mathbf{s} = \mathbf{0} \end{array} \tag{4.73}$$

The K-T condition of the problem 4.73 is

$$\begin{cases} \bigtriangledown^2 f(\mathbf{x})^T\delta_x + \bigtriangledown f(\mathbf{x}) + A^T\lambda = 0 \\ A\delta_x = 0 \end{cases} \tag{4.74}$$

We can rewrite (4.71) into matrix form, which is

$$\begin{bmatrix} \bigtriangledown^2 f(\mathbf{x}) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \lambda \end{bmatrix} = \begin{bmatrix} - \bigtriangledown f(\mathbf{x}) \\ 0 \end{bmatrix} \tag{4.75}$$

4.4.2.1 Termination condition

Define the Newton decrement as

$$\mathcal{K}(\mathbf{x}) = (\delta_x^T \bigtriangledown^2 f(\mathbf{x})\delta_x)^{\frac{1}{2}} \tag{4.76}$$

Since

$$\frac{\partial f(\mathbf{x}+\alpha\delta_x)}{\partial\alpha}|_{\alpha=0} = -\mathcal{K}(\mathbf{x})^2 \tag{4.77}$$

So the algorithm should terminate when $\mathcal{K}(\mathbf{x})$ is small.

4.4.2.2 Algorithm

Algorithm 9: Newton method with feasible start

Data: Cost function f, feasible region S
$x^{(0)} \in \mathbf{int}S$, $\epsilon > 0$, $k := 0$;
while $\mathcal{K}(\mathbf{x}^{(k)}) \geq \epsilon$ **do**
  Compute δ_x and $\mathcal{K}(\mathbf{x}^{(k)})$;
  Line search for step size α;
  $\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} + \alpha\delta_x$;
  $k := k + 1$;
end
return: $\mathbf{x}^{(k)}$

4.4.3 Newton method with infeasible start

Consider the case that initial point $\mathbf{x} \notin S$, we can apply

$$A\mathbf{x} = \mathbf{b} \tag{4.78}$$

to simplify (4.71). In this case, we can write the matrix form of the iteration of Newton step δ_x and λ by

$$\begin{bmatrix} \bigtriangledown^2 f(\mathbf{x}) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \lambda \end{bmatrix} = -\begin{bmatrix} \bigtriangledown f(\mathbf{x}) \\ A\mathbf{x} - \mathbf{b} \end{bmatrix} \tag{4.79}$$

4.4.3.1 A Primal-dual Interpretation

Recall problem 4.69, we can derive the K-T condition of the problem

$$\begin{cases} \bigtriangledown f(\mathbf{x}) + A^T\lambda = 0 \\ A\mathbf{x} = \mathbf{b} \end{cases} \tag{4.80}$$

Define $r : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n \times \mathbb{R}^p$ as

$$r(\mathbf{x}, \lambda) = (\bigtriangledown f(\mathbf{x}) + A^T\lambda, A\mathbf{x} - \mathbf{b})^T \tag{4.81}$$

where the first and second term is called the *dual* and *primal residual*, respectively. Then the K-T condition can be expressed as

$$r(\mathbf{x}, \lambda) = 0 \tag{4.82}$$

Apply Lagrange-Newton iteration to $r(\mathbf{x}, \lambda)$ we can derive

$$r(\mathbf{x} + \delta_x, \lambda + \delta_\lambda) \approx r(\mathbf{x}, \lambda) + J_r(\mathbf{x}, \lambda) \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = 0 \tag{4.83}$$

that is

$$\begin{bmatrix} \nabla^2 f(\mathbf{x}) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_\lambda \end{bmatrix} = - \begin{bmatrix} \nabla f(\mathbf{x}) + A^T \lambda \\ A\mathbf{x} - \mathbf{b} \end{bmatrix} \tag{4.84}$$

Notice that $\lambda := \lambda + \delta_\lambda$, then (4.79) is equivalent to (4.84).

4.4.3.2 Algorithm

Algorithm 10: Newton method with infeasible start

Data: Cost function f, feasible region S

$x^{(0)} \in \mathbf{dom} f$, $\epsilon > 0$, $\tau \in (0, 1/2)$, $\gamma \in (0, 1)$, $k := 0$;

while $A\mathbf{x}^{(k)} \neq \mathbf{b}$ *or* $\| r(\mathbf{x}^{(k)}, \lambda^{(k)}) \|_2 \geq \epsilon$ **do**

  Compute δ_x and δ_λ;

  Backtracking line search for step size α;

  $\alpha := 1$;

  while $\| r(\mathbf{x}^{(k)} + \alpha\delta_x, \lambda^{(k)} + \alpha\delta_\lambda) \|_2 > (1 - \tau\alpha) \| r(\mathbf{x}^{(k)}, \lambda^{(k)}) \|_2$ **do**

    $\alpha := \gamma\alpha$;

  end

  $\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} + \alpha\delta_x$;

  $\lambda^{(k+1)} := \lambda^{(k)} + \alpha\delta_\lambda$;

  $k := k + 1$;

end

return: $\mathbf{x}^{(k)}$

4.5 Interior point method

For inequality constrained convex problem

$$\begin{array}{ll} \min & f_0(\mathbf{x}) \\ s.t. & f_i(\mathbf{x}) \leq 0, i = 1, ..., m \\ & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.85}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}, i = 0, ..., m$ are convex and twice continuously differentiable, and $A \in \mathbb{R}^{p \times n}$ with $\mathbf{rank} A = p < n$. We assume that an optimal $\mathbf{x}^*$ exists and denote the optimal value $f_0(\mathbf{x}^*)$ as v^*.

We also assume that the problem is strictly feasible, i.e., $\exists \mathbf{x} \in \mathcal{D}$ satisfying $A\mathbf{x} = \mathbf{b}$ and $f_i(\mathbf{x}) < 0, i = 1, ..., m$.

This means that Slater's constraint qualification holds, and therefore strong duality holds, so there exists dual optimal $\lambda^* \in \mathbb{R}^m$, $\nu^* \in \mathbb{R}^p$, which together with $\mathbf{x}^*$ staisfy KKT conditions

$$\begin{array}{llll} \bigtriangledown f_0(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \bigtriangledown f_i(\mathbf{x}^*) & +A^T\nu^* & = 0 & \\ & \lambda^* & \geq 0 & \\ & f_i(\mathbf{x}^*) & \leq 0, & i = 1, ..., m \\ & A\mathbf{x}^* & = \mathbf{b} & \\ & \lambda_i^* f_i(\mathbf{x}^*) & = 0, & i = 1, ..., m \end{array} \tag{4.86}$$

4.5.1 Barrier interior-point method

Recall the barrier method in Chapter 3, we can rewrite inequality constrained problem 4.85 into

$$\begin{array}{ll} \min & f_0(\mathbf{x}) + \sum_{i=1}^m I_-(f_i(\mathbf{x})) \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.87}$$

where

$$I_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases} \tag{4.88}$$

In barrier method, we approximate the indicator function I_- by

$$\hat{I}_-(u) = -\frac{1}{t}\log(-u) \tag{4.89}$$

Obviously $\hat{I}_-$ is convex and differentiable. Then we substitude $\hat{I}_-$ for I_- in (4.87), result in

$$\begin{array}{ll} \min & f_0(\mathbf{x}) - \sum_{i=1}^m \frac{1}{t}\log(f_i(\mathbf{x})) \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.90}$$

The function

$$\Phi(\mathbf{x}) = -\sum_{i=1}^m \log(-f_i(\mathbf{x})) \tag{4.91}$$

is called the *logarithmic barrier* with

$$\mathbf{dom}\Phi = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) < 0, i = 1, ..., m\} \tag{4.92}$$

4.5.1.1 Central path

We rewrite problem 4.90 into an equivalent form

$$\begin{array}{ll} \min & tf_0(\mathbf{x}) + \Phi(\mathbf{x}) \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.93}$$

We assume that the problem 4.93 can be solved by Newton method and has unique solution for each $t > 0$.

For $t > 0$ we define $\mathbf{x}^*(t)$ as the solution of (4.93), the set of points $\mathbf{x}^*(t), t > 0$ us called *central path*.

From the K-T condition of (4.93), we have $\mathbf{x}^*(t)$ satisfies $\exists \nu \in \mathbb{R}^p$ such that

$$t \bigtriangledown f_0(\mathbf{x}^*(t)) + \bigtriangledown\Phi(\mathbf{x}^*(t)) + A^T\nu = 0 \tag{4.94}$$

$$t \bigtriangledown f_0(\mathbf{x}^*(t)) + \sum_{i=1}^{m} \frac{1}{-f_i(\mathbf{x}^*(t))} \bigtriangledown f_i(\mathbf{x}^*(t)) + A^T\nu = 0 \tag{4.95}$$

Define

$$\lambda_i^*(t) = -\frac{1}{tf_i(\mathbf{x}^*(t))}, i = 1, ..., m \tag{4.96}$$

$$\nu^*(t) = \nu/t \tag{4.97}$$

From $f_i(\mathbf{x}^*(t)) < 0, i = 1, ..., m$, we have $\lambda^*(t) > 0$.

Then (4.95) can be expressed as

$$\bigtriangledown f_0(\mathbf{x}^*(t)) + \sum_{i=1}^{m} \lambda^*(t) \bigtriangledown f_i(\mathbf{x}^*(t)) + A^T\nu^*(t) = 0 \tag{4.98}$$

We can see that $\mathbf{x}^*(t)$ is the minima of the Lagrangian

$$\mathcal{L}(\mathbf{x}, \lambda^*(t), \nu^*(t)) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^*(t) f_i(\mathbf{x}) + \nu^*(t)^T(A\mathbf{x} - \mathbf{b}) \tag{4.99}$$

due to the convexsity of (4.99). That is to say,

$$g(\lambda^*(t), \nu^*(t)) = \mathcal{L}(\mathbf{x}^*(t), \lambda^*(t), \nu^*(t) = \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda^*(t), \nu^*(t)) \tag{4.100}$$

Notice that

$$g(\lambda^*(t), \nu^*(t)) = f_0(\mathbf{x}^*(t)) + \sum_{i=1}^{m} \lambda_i^*(t) f_i(\mathbf{x}^*(t)) + \nu^*(t)^T(A\mathbf{x}^*(t) - \mathbf{b}) \tag{4.101}$$

$$= f_0(\mathbf{x}^*(t)) + \sum_{i=1}^{m} -\frac{1}{tf_i(\mathbf{x}^*(t))} f_i(\mathbf{x}^*(t)) \tag{4.102}$$

$$= f_0(\mathbf{x}^*(t)) - \frac{m}{t} \tag{4.103}$$

Then, we have

$$f_0(\mathbf{x}^*(t)) = g(\lambda^*(t), \nu^*(t)) + \frac{m}{t} \leq g(\lambda^*, \nu^*) + \frac{m}{t} = f_0(\mathbf{x}^*) + \frac{m}{t} \tag{4.104}$$

$$f_0(\mathbf{x}^*(t)) - f_0(\mathbf{x}^*) \leq \frac{m}{t} \tag{4.105}$$

that is, $\mathbf{x}^*(t) \to \mathbf{x}^*$.

4.5.1.2 Algorithm

Algorithm 11: Barrier interior-point algorithm

Data: Cost function f, feasible region S
Strictly feasible $\mathbf{x}^{(0)}$, $\mu > 1$, $\epsilon > 0$, $t > 0$, $k := 0$;
while $m/t \geq \epsilon$ **do**
  Use Newton method to compute $\mathbf{x}^{(k)} = \mathbf{x}^*(t)$ with initial point $\mathbf{x}^{(k-1)}$;
  $t := \mu t$;
  $k := k + 1$;
end
return: $\mathbf{x}^{(k)}$

4.5.1.3 Discussion

The step that use Newton method to compute $\mathbf{x}^*(t)$ requires iterations, which we called the *inner iterations*.

Selection of μ If μ is large, then $\mathbf{x}^{(k+1)}$ might be dramatically different from $\mathbf{x}^{(k)}$. That means that more inner iterations will be required.

If μ is small, then less inner iterations is required but on the contrary, there will be a large number of outer iterations.

The influence of the selection of t can be conducted similarly.

4.5.1.4 Newton step for computing $\mathbf{x}^*(t)$

For Step 1 in the Barrier interior-point method, i.e., solve $\mathbf{x}^*(t)$ in (4.93), we do not necessarily need the exact solution, an approximate $\mathbf{x}^*(t)$ is enough.

The Newton method for problem 4.93 is equivalent to solve the problem

$$\begin{array}{ll} \min & (\mathbf{x}+\mathbf{s})^T(t\bigtriangledown^2 f_0(\mathbf{x}) + \bigtriangledown^2\Phi(\mathbf{x}))(\mathbf{x}+\mathbf{s}) + (t\bigtriangledown f_0(\mathbf{x}) + \bigtriangledown\Phi(\mathbf{x}))^T(\mathbf{x}+\mathbf{s}) \\ s.t. & A(\mathbf{x}+\mathbf{s}) = \mathbf{b} \end{array} \tag{4.106}$$

Then we apply the Newton method for equality constrained convex optimization problems with feasible start, that is,

$$\begin{bmatrix} t\bigtriangledown^2 f_0(\mathbf{x}) + \bigtriangledown^2\Phi(\mathbf{x}) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \nu \end{bmatrix} = \begin{bmatrix} t\bigtriangledown f_0(\mathbf{x}) + \bigtriangledown\Phi(\mathbf{x}) \\ 0 \end{bmatrix} \tag{4.107}$$

The Newton step above can be interpreted as solving the *modified KKT equations*

$$\begin{array}{lll} \bigtriangledown f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i \bigtriangledown f_i(\mathbf{x}) & +A^T\nu & = 0 \\ & A\mathbf{x} & = \mathbf{b} \\ & -\lambda_i f_i(\mathbf{x}) & = \frac{1}{t}, \quad i = 1, ..., m \end{array} \tag{4.108}$$

when δ_x is small.

4.5.1.5 Basic phase I method

Notice that the Barrier interior-point method requires a strictly feasible starting point $\mathbf{x}^{(0)}$, so we need to use a algorithm to find it.

A strictly feasible point can be found by solving the following problem

$$\begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^n, s\in\mathbb{R}} & s \\ s.t. & f_i(\mathbf{x}) \leq s, \quad i = 1, .., m \\ & A\mathbf{x} = \mathbf{b} \end{array} \tag{4.109}$$

For every $\mathbf{x}$, we can find some proper s such that $(\mathbf{x}, s)$ is feasible, so we can apply the Barrier interior-point method to solve (4.109) to derive a strictly feasible point for the Barrier interior-point method.

4.5.2 Primal-dual interior-point method

5 Sparse Optimization

5.1 Compressed Sensing

5.1.1 Problem formulation

$$(\mathrm{P}_0) \quad \begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^n} & \| \mathbf{x} \|_0 \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{5.1}$$

The definition above means to find the sparsest solution for underdetermined linear equation $A\mathbf{x} = \mathbf{b}$ ($A \in \mathbb{R}^{m\times n}, m << n$).

Definition 5.1.1 (spark)**.** *The* spark *of a given matrix A is the smallest number of columns from A that are linearly dependent.*

Theorem 5.1.1. *If a system of linear euqations $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ obeying $\| \mathbf{x} \|_0 < \frac{spark(A)}{2}$, this solution is necessarily the sparsest possible.*

Definition 5.1.2. *The mutual coherence of a given matrix A is the largest absolute normalized inner product between different columns from A. Denoting the k-th column in A by $\mathbf{a}_k$, the mutual coherence is given by*

$$\mu(A) = \max_{1\le i\neq j\le n} \frac{|\mathbf{a}_i^T \mathbf{a}_j|}{\| \mathbf{a}_i \|_2 \| \mathbf{a}_j \|_2} \tag{5.2}$$

Lemma 5.1.2. *For any matrix $A \in \mathbb{R}^{m\times n}$, the following relationship holds:*

$$spark(A) \ge 1 + \frac{1}{\mu(A)} \tag{5.3}$$

Then we have the following theorem:

Theorem 5.1.3. *If a system of linear euqations $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ obeying $\| \mathbf{x} \|_0 < (1 + \frac{1}{\mu(A)})/2$, this solution is necessarily the sparsest possible.*

5.1.2 Restricted Isometry Property

Consider problem

$$(\mathrm{P}_1) \quad \begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^n} & \| \mathbf{x} \|_1 \\ s.t. & A\mathbf{x} = \mathbf{b} \end{array} \tag{5.4}$$

5.1.3 Pursuit Algorithms

5.1.3.1 Orthogonal Matching Pursuit

Algorithm 12: OMP Algorithm

Data:
;
while do
end
return:

5.1.3.2 Basis Pursuit