Optimization Algorithm Notes

Depu Meng

Feb. 2019

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 Topics Involved

- Unconstrained optimization
- Constrained optimization
- Convex optimization
- Sparse optimization
- Stochastic optimization
- Combinational optimization
- Global optimization

1.3 Basic Concepts

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

$$\min f(x) \tag{1}$$
$$s.t. x \in \mathcal{S} \subset \mathbb{R}^n \tag{2}$$

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

We often denote constrained optimization Problem as

$$\min f(x) \tag{3}$$
$$s.t. \quad g_i(x) \geq 0, i = 1, ..., n \tag{4}$$
$$b_i(x) = 0, i \in 1, ..., m \tag{5}$$

Definition 1. *Global Optimality. For global optimal value $x^* \in \mathcal{S}$,*

$$f(x^*) \leq f(x), \forall x \in \mathcal{S} \tag{6}$$

Definition 2. *Local Optimality. For local optimal value $x^* \in \mathcal{S}$, $\exists U(x^*)$, such that*

$$f(x^*) \leq f(x), \forall x \in \mathcal{S} \cap U(x^*) \tag{7}$$

Definition 3. *Feasible direction. Let $x \in \mathcal{S}$, $d \in \mathbb{R}^n$ is a non-zero vector. if $\exists \delta > 0$, such that*

$$x + \lambda d \in \mathcal{S}, \forall \lambda \in (0, \delta) \tag{8}$$

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

Definition 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*

$$f(x + \lambda d) < f(x), \forall \lambda \in (0, \delta) \tag{9}$$

*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.*

1.4 Optimal Conditions

Unconstrained Optimization

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

$$\bigtriangledown f(x) = 0 \tag{10}$$

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

$$\bigtriangledown f(x) = 0 \tag{11}$$

$$\bigtriangledown^2 f(x) \geq 0 \tag{12}$$

Constrained Optimization

Theorem 1. *Fritz-John Condition*

For constrained optimization problem

$$\min f(x) \tag{13}$$

$$s.t. \quad g_i(x) \geq 0, i = 1, ..., n \tag{14}$$

$$h_i(x) = 0, i \in 1, ..., m \tag{15}$$

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

$$\lambda_0 \bigtriangledown f(x) - \sum_{i \in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^{m} \mu_j \bigtriangledown h_j(x) = 0 \tag{16}$$

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

$$\sum_{j=1}^{m} \bigtriangledown \mu_j h_j(x) = 0 \tag{17}$$

Let $\lambda_0, \lambda_i, i \in I(x) = 0$, then (13) holds.

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

$$F_g = F(x,g) = \{d| \bigtriangledown g_i(x)^T d > 0, i \in I(x)\} \tag{18}$$

$$F_h = F(x,h) = \{d| \bigtriangledown h_j(x)^T d = 0, j = 1, ..., m\} \tag{19}$$

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

$$F_g \cap F_h \cap D(x, f) = \varnothing \tag{20}$$

that is

$$\begin{cases} \bigtriangledown f(x)^T d < 0 \\ \bigtriangledown g_i(x)^T d > 0, i \in I(x) \\ \bigtriangledown h_j(x)^T d = 0, j = 1, ..., m \end{cases} \tag{21}$$

has no solution. Let

$$A = \{\bigtriangledown f(x)^T, - \bigtriangledown g_i(x)\}^T, i \in I(x) \tag{22}$$

$$B = \{- \bigtriangledown h_j(x)\}, j = 1, ..., m \tag{23}$$

Then (21) is equivalent to

$$\begin{cases} A^T d < 0 \\ B^T d = 0 \end{cases} \tag{24}$$

has no solution.
Denote

$$S_1 = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} | y_1 = A^T d, y_2 = B^T d, d \in \mathbb{R}^n \right\} \tag{25}$$

$$S_2 = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} | y_1 < 0, y_2 = 0 \right\} \tag{26}$$

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

$$p_1^T A^T d + p_2^T B^T d \geq p_1^T y_1 + p_2^T y_2, \forall d \in \mathbb{R}^n, \forall \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in CL(S_2) \tag{27}$$

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

$$p_1 \geq 0 \tag{28}$$

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

$$(p_1^T A^T + p_2^T B^T)d \geq 0 \tag{29}$$

$$(Ap_1 + Bp_2)^T d \geq 0 \tag{30}$$

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

$$Ap_1 + Bp_2 = 0 \tag{31}$$

From above, we have

$$\begin{cases} Ap_1 + Bp_2 = 0 \\ p_1 \geq 0 \end{cases} \tag{32}$$

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

$$\begin{cases} \lambda_0 \bigtriangledown f(x) - \sum_{i \in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^m \mu_j \bigtriangledown h_j(x) = 0 \\ \lambda_i \geq 0 \end{cases} \tag{33}$$

Theorem 2. *Kuhn-Tucker Condition*
For constrained optimization problem

$$\min f(x) \tag{34}$$

$$s.t. \quad g_i(x) \geq 0, i = 1, ..., n \tag{35}$$

$$h_i(x) = 0, i \in 1, ..., m \tag{36}$$

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*

$$\bigtriangledown f(x) - \sum_{i \in I(x)} \lambda_i \bigtriangledown g_i(x) - \sum_{j=1}^m \mu_j \bigtriangledown h_j(x) = 0 \tag{37}$$

1.5 Descent function

Definition 5. *Descent function. Denote solution set* $\Omega \in X$, $\mathcal{A}$ *is an algorithm on* X, $\psi : X \to \mathbb{R}$. *If*

$$\psi(y) < \psi(x), \quad \forall x \notin \Omega, y \in \mathcal{A}(x) \tag{38}$$

$$\psi(y) \leq \psi(x), \quad \forall x \in \Omega, y \in \mathcal{A}(x) \tag{39}$$

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

1.6 Convergence of Algorithm

Theorem 3. *$\mathcal{A}$ 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.

1.7 Search Methods

Line Search

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

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \tag{40}$$

. search α_k in 1-D space.

Trust Region

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

$$s^{(k)} = \arg\min Q_k(s) \tag{41}$$

$$x^{(k+1)} = x^{(k)} + s^{(k)} \tag{42}$$

2 Unconstrained Optimization

2.1 Gradient Based Methods

$$\min_{x \in \mathbb{R}^n} f(x) \tag{43}$$

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

2.2 Determine Search Direction

First-order gradient method

For unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} f(x) \tag{44}$$

We have

$$f(x) = f(x^{(k)}) + \bigtriangledown f(x^{(k)})^T (x - x^{(k)}) + O(\| x - x^{(k)} \|^2) \tag{45}$$

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

$$f(x^{(k)} + \alpha_k d^{(k)}) < f(x^{(k)}) \tag{46}$$

Second-order gradient method – Newton Direction

$$f(x) = f(x^{(k)}) + \bigtriangledown f(x^{(k)})^T (x - x^{(k)}) \tag{47}$$

$$+ \frac{1}{2} (x - x^{(k)})^T \bigtriangledown^2 f(x^{(k)}) (x - x^{(k)}) + O(\| x - x^{(k)} \|^3) \tag{48}$$

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 $x^{(k)}$.

2.3 Determine Step Factor – Line Search

$$\min_{\alpha \geq 0} \varphi(\alpha) = f(x^{(k)} + \alpha d^{(k)}) \tag{49}$$

Exact Line Search

Solve Line Search problem in finite iterations.

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

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{50}$$

$$\varphi(\alpha) \geq \varphi(0) + (1-\rho)\alpha\varphi'(0) \tag{51}$$

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

$$\varphi(\alpha) \leq \varphi(0) + \rho\alpha\varphi'(0) \tag{52}$$

$$\varphi'(\alpha) \geq \sigma\varphi'(0) \tag{53}$$

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

2.4 Global Convergence

Theorem 4. *Assume $\bigtriangledown f(x)$ exists and uniformly continuous on level set $L(x^{(0)}) = \{x|f(x) \leq f(x^{(0)})\}$. Denote $\theta^{(k)}$ as the angle between $d^{(k)}$ and $-\bigtriangledown f(x^{(k)})$.*

$$\theta^{(k)} \leq \frac{\pi}{2} - \mu \tag{54}$$

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.

2.5 Steepest Descent Method

Steepest Descent Method is a Line Search Method.

$$x^{(k+1)} = x^{(k)} - \alpha_k \bigtriangledown f(x^{(k)}) \tag{55}$$

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.

2.6 Newton Method

Newton Method is also a Line Search Method.

$$f(x^{(k)} + s) \approx q^{(k)}(s) f(x^{(k)}) + g^{(k)^T} s + \frac{1}{2} s^T G_k s \tag{56}$$

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

$$s = G_k^{-1} g^{(k)} \tag{57}$$

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

Analysis on quadratic function

For positive definite quadratic function

$$f(x) = \frac{1}{2} x^T G x - c^T x \tag{58}$$

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

$$\begin{align} d^{(0)} &= H_0 \bigtriangledown f(x^{(0)}) \tag{59} \\ &= G^{-1}(G x^{(0)} - c) \tag{60} \\ &= x^{(0)} - G^{-1} c \tag{61} \\ &= x^{(0)} - x^* \tag{62} \end{align}$$

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

For general non-linear functions, if we follow

$$x^{(k+1)} = x^{(k)} - G_k^{-1} g^{(k)} \tag{63}$$

we called it Newton Method.

Convergence Rate of Newton Method

Theorem 5. *$f \in \mathcal{C}^2$, $x^{(k)}$ 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),*

$$|G_{ij}(x) - G_{ij}(y)| \leq \beta \parallel x - y \parallel \tag{64}$$

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

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

$$g(x - h) = g(x) - G(x)h + O(\parallel h \parallel^2) \tag{65}$$

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

$$g(x^*) = g(x^{(k)}) - G(x^{(k)})(h^{(k)}) + O(\parallel h^{(k)} \parallel^2) = 0 \tag{66}$$

From Lipschitz Condition, we can easily get $G(x^{(k)})^{-1}$ is finite. Then we left multiply $G(x^{(k)})^{-1}$ to Equation (66)

$$0 = G(x^{(k)})^{-1} g(x^{(k)}) - h^{(k)} + O(\parallel h^{(k)} \parallel^2) \tag{67}$$

$$= x^* - x^{(k)} + G(x^{(k)})^{-1} g(x^{(k)}) + O(\parallel h^{(k)} \parallel^2) \tag{68}$$

$$= x^* - x^{(k+1)} + O(\parallel h^{(k)} \parallel^2) \tag{69}$$

$$= -h^{(k+1)} + O(\parallel h^{(k)} \parallel^2) \tag{70}$$

i.e.,

$$\parallel h^{(k+1)} \parallel = O(\parallel h^{(k)} \parallel^2) \tag{71}$$

2.7 Quasi-Newton Methods

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

$$\bigtriangledown^2 f(x^{(k)}) s^{(k)} \approx y^{(k)} \tag{72}$$

or

$$\bigtriangledown^2 f(x^{(k)})^{-1} y^{(k)} \approx s^{(k)} \tag{73}$$

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

$$H_{k+1} y^{(k)} \approx s^{(k)} \tag{74}$$

or

$$y^{(k)} \approx B_{k+1} s^{(k)} \tag{75}$$

we called (74), (75) *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

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

Assume

$$H_{k+1} = H_k + a\mathbf{u}\mathbf{u}^T, \quad a \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^n \tag{76}$$

From the Quasi-Newton Conditions, we have

$$H_{k+1}\mathbf{y}^{(k)} = \mathbf{s}^{(k)} \tag{77}$$

$$H_k\mathbf{y}^{(k)} + a\mathbf{u}\mathbf{u}^T\mathbf{y}^{(k)} = \mathbf{s}^{(k)} \tag{78}$$

$$H_k\mathbf{y}^{(k)} + a\mathbf{u}^T\mathbf{y}^{(k)}\mathbf{u} = \mathbf{s}^{(k)} \tag{79}$$

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

$$H_{k+1} = \frac{(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})^T}{(\mathbf{s}^{(k)} - H_k\mathbf{y}^{(k)})^T\mathbf{y}^{(k)}} \tag{80}$$

(79) 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

$$H_{k+1} = H_k + a\mathbf{u}\mathbf{u}^T + b\mathbf{v}\mathbf{v}^T, \quad a, b \in \mathbb{R}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \tag{81}$$

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

$$\begin{cases} \mathbf{u} = \mathbf{s}^{(k)}, & a\mathbf{u}^T\mathbf{y} = 1 \\ \mathbf{v} = H_k\mathbf{y}^{(k)}, & b\mathbf{v}^T\mathbf{y} = -1 \end{cases} \tag{82}$$

So that we have

$$H_{k+1} = H_k + \frac{\mathbf{s}^{(k)}\mathbf{s}^{(k)T}}{\mathbf{s}^{(k)T}\mathbf{y}^{(k)}} - \frac{H_k\mathbf{y}^{(k)}\mathbf{y}^{(k)T}H_k}{\mathbf{y}^{(k)T}H_k\mathbf{y}^{(k)}} \tag{83}$$

We called (83) the DFP (Davidon-Fletcher-Powell) update.

From Quasi-Newton Condition (75), we can get the BFGS (Broyden-Fletcher-Goldfarb-Shanno) update

$$B_{k+1}^{(BFGS)} = B_k + \frac{\mathbf{y}^{(k)}\mathbf{y}^{(k)T}}{\mathbf{y}^{(k)T}\mathbf{s}^{(k)}} - \frac{B_k\mathbf{s}^{(k)}\mathbf{s}^{(k)T}B_k}{\mathbf{s}^{(k)T}B_k\mathbf{s}^{(k)}} \tag{84}$$

Inverse of SR1 update

Theorem 6 (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*

$$(A + a\mathbf{u}\mathbf{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^T A^{-1}}{1 + \mathbf{v}^T A^{-1}\mathbf{u}} \tag{85}$$

2.8 Conjugate Gradient Method

Definition 6. *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 1. *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 6, we have

$$\mathbf{d}^{(i)T}G\mathbf{d}^{(j)} = 0, \forall i, j, i \neq j \tag{86}$$

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

$$\mathbf{d}^{(t)} = \sum_{j=0}^{k} c_j \mathbf{d}^{(j)} \tag{87}$$

then

$$\mathbf{d}^{(t)T}G\mathbf{d}^{(i)} = \sum_{j=0}^{k} c_j \mathbf{d}^{(j)}G\mathbf{d}^{(i)} = c_i\mathbf{d}^{(i)}G\mathbf{d}^{(i)} \neq 0 \tag{88}$$

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 7 (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*

$$\mathcal{V} = \{\mathbf{x} | \mathbf{x} = \mathbf{x}^{(0)} + \sum_{j=0}^{k} \beta_j \mathbf{d}^{(j)}, \forall \beta_j \in \mathbb{R}\} \tag{89}$$

Proof. Firstly, from Lemma 2, we have $\{\mathbf{d}^{(0)}, ..., \mathbf{d}^{(k)}\}$ are linearly independent. So we only need to prove that for all $k < n$

$$\mathbf{g}^{(k+1)T}\mathbf{d}^{(j)} = 0, j = 0, ..., k \tag{90}$$

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

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

$$\mathbf{g}^{(j+1)T}\mathbf{d}^{(j)} = 0 \tag{91}$$

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

Notice that

$$\mathbf{g}^{(k+1)} - \mathbf{g}^{(k)} = G(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \alpha_k G \mathbf{d}^{(k)} \tag{92}$$

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

$$\mathbf{g}^{(k+1)T}\mathbf{d}^{(j)} = (\sum_{m=j+1}^{k} (\mathbf{g}^{(m+1)T} - \mathbf{g}^{(m)T}) + \mathbf{g}^{(j+1)T})\mathbf{d}^{(j)} \tag{93}$$

$$= \sum_{m=j+1} \alpha_m \mathbf{d}^{(m)T} G \mathbf{d}^{(j)} + \mathbf{g}^{(j+1)T}\mathbf{d}^{(j)} \tag{94}$$

$$= 0 \tag{95}$$

□

Lemma 2. *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*

$$\mathbf{g}^{(k)T}\mathbf{g}^{(j)} = 0, \forall j = 0, ..., k-1 \tag{96}$$

Proof. From Theorem 7, we have

$$\mathbf{g}^{(k)T}\mathbf{g}^{(j)} = \mathbf{g}^{(k)T}(-\mathbf{d}^{(j)} + \sum_{i=0}^{j-1}\beta_i^{(j)}\mathbf{d}^{(i)}) = 0 \tag{97}$$

□

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.

$$\mathbf{g}(\mathbf{x}) = G\mathbf{x} + \mathbf{c} \tag{98}$$

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

$$\beta_0^{(1)} = \frac{\mathbf{g}^{(1)T}\mathbf{g}^{(1)}}{\mathbf{g}^{(0)T}\mathbf{g}^{(0)}} \tag{99}$$

Proof. From (92), we have

$$\mathbf{d}^{(1)T}G\mathbf{d}^{(0)} = 0 \tag{100}$$

$$\Leftrightarrow \quad \mathbf{d}^{(1)T}(\mathbf{g}^{(1)} - \mathbf{g}^{(0)}) = 0 \tag{101}$$

$$\Leftrightarrow \quad (\mathbf{g}^{(1)} + \beta_0^{(1)}\mathbf{g}^{(0)})^T(\mathbf{g}^{(1)} - \mathbf{g}^{(0)}) = 0 \tag{102}$$

$$\Leftrightarrow \quad \mathbf{g}^{(1)T}\mathbf{g}^{(1)} - \beta_0^{(1)}\mathbf{g}^{(0)T}\mathbf{g}^{(0)} = 0 \tag{103}$$

$$\Leftrightarrow \quad \beta_0^{(1)} = \frac{\mathbf{g}^{(1)T}\mathbf{g}^{(1)}}{\mathbf{g}^{(0)T}\mathbf{g}^{(0)}} \tag{104}$$

□

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

$$\mathbf{d}^{(k)T}G\mathbf{d}^{(j)} = 0 \tag{105}$$

$$(-\mathbf{g}^{(k)T} + \sum_{i=0}^{k-1}\beta_i^{(k)}\mathbf{d}^{(i)T})G\mathbf{d}^{(j)} = 0 \tag{106}$$

$$-\mathbf{g}^{(k)T}G\mathbf{d}^{(j)} + \beta_j^{(k)}\mathbf{d}^{(j)T}G\mathbf{d}^{(j)} = 0 \tag{107}$$

so we have

$$\beta_j^{(k)} = \frac{\mathbf{g}^{(k)T} G \mathbf{d}^{(j)}}{\mathbf{d}^{(j)T} G \mathbf{d}^{(j)}} = \frac{\mathbf{g}^{(k)T}(\mathbf{g}^{(j+1)} - \mathbf{g}^{(j)})}{\mathbf{d}^{(j)T}(\mathbf{g}^{(j+1)} - \mathbf{g}^{(j)})} \tag{108}$$

From Lemma 2, we have

$$\mathbf{g}^{(k)T} \mathbf{g}^{(j)} = 0, \forall j = 0, ..., k-1 \tag{109}$$

So

$$\beta_j^{(k)} = 0, j = 0, ..., k-2 \tag{110}$$

$$\beta_{k-1}^{(k)} = \frac{\mathbf{g}^{(k)T}(\mathbf{g}^{(k)}}{\mathbf{g}^{(k-1)T}(\mathbf{g}^{(k-1)}} \tag{111}$$

2.9 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 7 (Trust Region).

$$\Omega_k = \{\mathbf{x} \in \mathbb{R}^n \mid \ \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq e_k\} \tag{112}$$

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 $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$.

Trust Region Subproblem

$$\min_{\|\mathbf{s}\| \leq e_k} q^{(k)}(\mathbf{s}) = f(\mathbf{x}^{(k)}) + \mathbf{g}^{(k)T}\mathbf{s} + \frac{1}{2}\mathbf{s}^T B_k \mathbf{s} \tag{113}$$

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.

How to select e_k

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

$$\mathrm{Act}_k = f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k)} + \mathbf{s}^{(k)}) \tag{114}$$

$$\mathrm{Pre}_k = q^{(k)}(\mathbf{0}) - q^{(k)}(\mathbf{s}^{(k)}) \tag{115}$$

Define

$$r_k = \frac{\mathrm{Act}_k}{\mathrm{Pre}_k} = \frac{f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k)} + \mathbf{s}^{(k)})}{q^{(k)}(\mathbf{0}) - q^{(k)}(\mathbf{s}^{(k)})} \tag{116}$$

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

$$e_{k+1} = \begin{cases} \eta_1 e_k & \text{if } r_k < \gamma_1 \\ e_k & \text{if } \gamma_1 < r_k < \gamma_2 \\ \min(\eta_2 e_k, \bar{e}) & \text{if } r_k \geq \gamma_2 \end{cases} \tag{117}$$

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

$$\min \ Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x}) \tag{118}$$

$$s.t. \ \ \mathbf{a}_i^T\mathbf{x} = b_i, i \in \mathcal{E} = \{1, ..., m_e\} \tag{119}$$

$$\mathbf{a}_i^T\mathbf{x} \geq b_i, i \in \mathcal{I} = \{m_e + 1, ..., m\} \tag{120}$$

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

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.

Equality Constrained Quadratic Programming

$$\min \ Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \tag{121}$$

$$s.t. \ \ A\mathbf{x} = \mathbf{b} \tag{122}$$

General Quadratic Programming

$$\min \ Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x}) \tag{123}$$

$$s.t. \ \ \mathbf{a}_i^T\mathbf{x} = b_i, i \in \mathcal{E} = \{1, ..., m_e\} \tag{124}$$

$$\mathbf{a}_i^T\mathbf{x} \geq b_i, i \in \mathcal{I} = \{m_e + 1, ..., m\} \tag{125}$$

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 8 (Active Set). *Denote* $\mathbf{x}^*$ *as a local minima of general quadratic problem* (123)*, then* $\mathbf{x}^*$ *must be a local minima of the equality constrained problem*

$$(\mathrm{EQ}) \begin{cases} \min \ \ Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ s.t. \ \ \mathbf{a}_i^T\mathbf{x} \ \ = b_i, i \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*) \end{cases} \tag{126}$$

Meanwhile, if $\mathbf{x}^*$ *is a feasible point of* (123)*, 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* (123)*.*

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.

$$\bigtriangledown Q(\mathbf{x}^*) - \sum_{i \in I(\mathbf{x}^*)} \lambda_i \mathbf{a}_i - \sum_{j \in \mathcal{E}} \mu_j \mathbf{a}_j = 0 \tag{127}$$

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

$$\bigtriangledown Q(\mathbf{x}^*) - \sum_{j \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)} \lambda_j \mathbf{a}_j = 0 \tag{128}$$

Appearently If $\mathbf{x}^*$ satisfies (127), then it also satisfies (128). On the other hand, if $\mathbf{x}^*$ satisfies (128) and $\lambda_i \geq 0, i \in I(\mathbf{x}^*)$, we have

$$\bigtriangledown Q(\mathbf{x}^*) - \sum_{j \in \mathcal{E} \cup \mathcal{I}(\mathbf{x}^*)} \lambda_j \mathbf{a}_j = 0 \tag{129}$$

$$\Leftrightarrow \bigtriangledown Q(\mathbf{x}^*) - \sum_{i \in I(\mathbf{x}^*)} \lambda_i \mathbf{a}_i - \sum_{j \in \mathcal{E}} \lambda_j \mathbf{a}_j = 0 \tag{130}$$

i.e., $\mathbf{x}^*$ satisfies (127). □

3.2 Non-linear Constrained Optimization

Equality Constrained Problem

Lagrange-Newton method

$$\min f(\mathbf{x}) \tag{131}$$

$$s.t.\ \mathbf{c}(\mathbf{x}) = \mathbf{0} \tag{132}$$

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

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

$$\bigtriangledown f(\mathbf{x}) - A(\mathbf{x})^T \lambda = \mathbf{0} \tag{133}$$

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

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

$$\begin{pmatrix} W(\mathbf{x}, \lambda) & -A(\mathbf{x})^T \\ -A(\mathbf{x}) & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = - \begin{pmatrix} \bigtriangledown f(\mathbf{x}) - A(\mathbf{x})^T \lambda \\ \mathbf{c}(\mathbf{x}) \end{pmatrix} \tag{134}$$

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

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

Here we can define

$$\psi(\mathbf{x}, \lambda) = \parallel \bigtriangledown f(\mathbf{x}) - A(\mathbf{x})^T \lambda \parallel^2 + \parallel \mathbf{c}(\mathbf{x}) \parallel^2 \tag{135}$$

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

$$\bigtriangledown \psi(\mathbf{x}, \lambda)^T \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = -2\psi(\mathbf{x}, \lambda) \neq 0 \tag{136}$$

Sequential Quadratic Programming method

(134) can be rewritten into

$$\begin{cases} W(\mathbf{x},\lambda)\delta_x + \bigtriangledown f(\mathbf{x}) = A(\mathbf{x})^T(\lambda+\delta_\lambda) \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\delta_x = \mathbf{0} \end{cases} \tag{137}$$

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

$$\min \frac{1}{2}\mathbf{d}^T W(\mathbf{x},\lambda)\mathbf{d} + \bigtriangledown f(\mathbf{x})^T\mathbf{d} \tag{138}$$

$$s.t.\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} = 0 \tag{139}$$

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

General Nonlinear Constrained Problem

Sequential Quadratic Programming method

$$\min f(\mathbf{x}) \tag{140}$$

$$s.t.\ c_i(\mathbf{x}) = 0, \quad i \in \mathcal{E} = \{1,...,m_e\} \tag{141}$$

$$c_i(\mathbf{x}) \geq 0, \quad i \in \mathcal{I} = \{m_e+1,...,m\} \tag{142}$$

Similarly, we can construct subproblem

$$\min \frac{1}{2}\mathbf{d}^T W\mathbf{d} + \mathbf{g}^T\mathbf{d} \tag{143}$$

$$s.t.\ c_i(\mathbf{x}) + \mathbf{a}_i(\mathbf{x})^T\mathbf{d} = 0, i \in \mathcal{E} \tag{144}$$

$$c_i(\mathbf{x}) + \mathbf{a}_i(\mathbf{x})^T\mathbf{d} \geq 0, i \in \mathcal{I} \tag{145}$$

Here, W is the Hesse matrix (or its approximation) of the Lagrange function of (140), $\mathbf{g} = \bigtriangledown f(\mathbf{x})$, $A(\mathbf{x}) = (\mathbf{a}_1(\mathbf{x}),...,\mathbf{a}_m(\mathbf{x})$.

Denote the solution to subproblem (143) as $\mathbf{d}$, the corresponding Lagrange multiplier vector $\bar{\lambda}$, so we have

$$\begin{cases} W\mathbf{d} + \mathbf{g} = A(\mathbf{x})^T\bar{\lambda} \\ \bar{\lambda}_i \geq 0, i \in \mathcal{I} \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} = 0, i \in \mathcal{E} \\ \mathbf{c}(\mathbf{x}) + A(\mathbf{x})\mathbf{d} \geq 0, i \in \mathcal{I} \end{cases} \tag{146}$$

Penalty method

For nonlinear constrained porblem (140), we can use objective function $f(\mathbf{x})$ and constraint function $\mathbf{c}(\mathbf{x})$ to construct *Penalty function*

$$P(\mathbf{x}) = P(f(\mathbf{x}), \mathbf{c}(\mathbf{x})) \tag{147}$$

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 $\mathbf{c}(\mathbf{x})_-$

$$\begin{cases} c_i(\mathbf{x})_- = c_i(\mathbf{x}), & i \in \mathcal{E} \\ c_i(\mathbf{x})_- = |\min\{0, c_i(\mathbf{x})\}|, & i \in \mathcal{I} \end{cases} \tag{148}$$

Consider simple penalty function

$$P_\sigma(\mathbf{x}) = f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{149}$$

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

Lemma 3 (Penalty method). *If $\mathbf{x}(\sigma)$ is a feasible point of nonlinear constrained problem* (140)*, then $\mathbf{x}(\sigma)$ aslo is the solution to* (140)*.*

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.,

$$P(\mathbf{x}(\sigma)) \leq P(\mathbf{x}_0),\ \forall \mathbf{x}_0 \in \mathbb{R}^n \tag{150}$$

$$f(\mathbf{x}(\sigma)) \leq f(\mathbf{x}_0),\ \forall \mathbf{x}_0 \in \mathcal{S} \tag{151}$$

that is, $\mathbf{x}(\sigma)$ is the solution to (140). □

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

Proof.

$$P_{\sigma_{k+1}}(\mathbf{x}(\sigma_{k+1})) = f(\mathbf{x}(\sigma_{k+1})) + \sigma_{k+1} \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 \tag{152}$$

$$\geq f(\mathbf{x}(\sigma_{k+1})) + \sigma_k \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 \tag{153}$$

$$\geq \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \sigma_k \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 \tag{154}$$

$$= P_{\sigma_k}(\mathbf{x}(\sigma_k)) \tag{155}$$

From the definition, we have

$$\begin{aligned} &f(\mathbf{x}(\sigma_k)) + \sigma_{k+1} \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2 && (156)\\ \geq &f(\mathbf{x}(\sigma_{k+1})) + \sigma_{k+1} \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 && (157)\\ \geq &f(\mathbf{x}(\sigma_{k+1})) + \sigma_{k} \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 && (158)\\ \geq &f(\mathbf{x}(\sigma_k)) + \sigma_{k} \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2 && (159) \end{aligned}$$

From the inequalities above, we have

$$\begin{aligned} &\sigma_k(\parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 - \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2) && (160)\\ \leq &f(\mathbf{x}(\sigma_{k+1})) - f(\mathbf{x}(\sigma_k)) && (161)\\ \leq &\sigma_{k+1}(\parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2 - \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2) && (162) \end{aligned}$$

So that

$$\parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel \geq \parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel \tag{163}$$

Then

$$0 \leq \sigma_k(\parallel \mathbf{c}(\mathbf{x}(\sigma_{k+1}))_- \parallel^2 - \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2) \leq f(\mathbf{x}(\sigma_{k+1})) - f(\mathbf{x}(\sigma_k)) \tag{164}$$

i.e.,

$$f(\mathbf{x}(\sigma_{k+1})) \geq f(\mathbf{x}(\sigma_k)) \tag{165}$$

□

Lemma 5. *Denote $\bar{\mathbf{x}}$ as the solution to problem* (140)*, then for all $\sigma_k > 0$,*

$$f(\bar{\mathbf{x}}) \geq P_{\sigma_k}(\mathbf{x}(\sigma_k)) \geq f(\mathbf{x}(\sigma_k)) \tag{166}$$

Proof. For all $\sigma_k > 0$,

$$\begin{aligned} f(\bar{\mathbf{x}}) &= \min_{\mathbf{x}\in\mathbb{R}^n} \lim_{\sigma\to\infty} f(\mathbf{x}) + \sigma \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 && (167)\\ &\geq \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \sigma_k \parallel \mathbf{c}(\mathbf{x})_- \parallel^2 && (168)\\ &= f(\mathbf{x}(\sigma_k)) + \sigma_k \parallel \mathbf{c}(\mathbf{x}(\sigma_k))_- \parallel^2 && (169)\\ &\geq f(\mathbf{x}(\sigma_k)) && (170) \end{aligned}$$

□

Lemma 6. *Let $\delta = \parallel \mathbf{c}(\mathbf{x}(\sigma))_- \parallel$, then $\mathbf{x}(\sigma)$ is also the solution to the problem*

$$\begin{aligned} &\min\ f(\mathbf{x}) && (171)\\ &s.t.\ \parallel \mathbf{c}(\mathbf{x})_- \parallel \leq \delta && (172) \end{aligned}$$

Proof. The problem is equivalent to

$$\min \ f(\mathbf{x}) \tag{173}$$
$$s.t. \ \ \| \mathbf{c}(\mathbf{x})_- \| \leq \| \mathbf{c}(\mathbf{x}(\sigma))_- \| \tag{174}$$

$$f(\mathbf{x}(\sigma)) + \sigma \| \mathbf{c}(\mathbf{x}(\sigma))_- \|^2 = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) + \sigma \| \mathbf{c}(\mathbf{x})_- \|^2 \tag{175}$$

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

$$f(\mathbf{x}(\sigma)) + \sigma \| \mathbf{c}(\mathbf{x}(\sigma))_- \|^2 \leq f(\mathbf{x}) + \sigma \| \mathbf{c}(\mathbf{x})_- \|^2 \tag{176}$$
$$f(\mathbf{x}(\sigma)) - f(\mathbf{x}) \leq \sigma(\| \mathbf{c}(\mathbf{x})_- \|^2 - \| \mathbf{c}(\mathbf{x}(\sigma))_- \|^2) \tag{177}$$

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

$$f(\mathbf{x}(\sigma)) - f(\mathbf{x}) \leq \sigma(\| \mathbf{c}(\mathbf{x})_- \|^2 - \| \mathbf{c}(\mathbf{x}(\sigma))_- \|^2) \leq 0 \tag{178}$$

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

4 Convex Optimization

4.1 Convex set

4.2 Convex function

4.3 Convex optimization

Quadratically constrained quadratic program

$$\min \frac{1}{2}\mathbf{x}^T P_0 \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \tag{179}$$

$$s.t. \frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 1, ..., m \tag{180}$$

$$A\mathbf{x} = \mathbf{b} \tag{181}$$

Second-order cone program

$$\min \mathbf{f}^T \mathbf{x} \tag{182}$$

$$s.t. \parallel A_i \mathbf{x} + \mathbf{b}_i \parallel \leq \mathbf{c}_i^T \mathbf{x} + \mathbf{d}_i, \quad i = 1, ..., m \tag{183}$$

$$F\mathbf{x} = \mathbf{g} \tag{184}$$

Lemma 7. *Any QCQP problem can be formulated as a SOCP problem.*

Proof. The QCQP problem is equivalent to

$$\min \ -r_0 \tag{185}$$

$$s.t. \frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 0, ..., m \tag{186}$$

$$A\mathbf{x} = \mathbf{b} \tag{187}$$

Then we need to prove that (121) can be formulated as (118).

$$\frac{1}{2}\mathbf{x}^T P_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0 \tag{188}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + 2(\mathbf{q}_i^T \mathbf{x} + r_i) \leq 0 \tag{189}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + 2(\mathbf{q}_i^T \mathbf{x} + r_i) + (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{190}$$

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + (\mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{191}$$

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

$$\Leftrightarrow \mathbf{x}^T P_i \mathbf{x} + (\mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2})^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{192}$$

$$\Leftrightarrow \ \parallel A_i \mathbf{x} \parallel^2 + \parallel \mathbf{q}_i^T \mathbf{x} + r_i + \frac{1}{2} \parallel^2 \leq (\mathbf{q}_i^T \mathbf{x} + r_i - \frac{1}{2})^2 \tag{193}$$

Let

$$A_i' = \begin{pmatrix} A \\ \mathbf{q}^T \end{pmatrix} \tag{194}$$

$$\mathbf{b}_i = \begin{pmatrix} \mathbf{0}_{n\times 1} \\ r_i + \frac{1}{2} \end{pmatrix} \tag{195}$$

From (123) and $\mathbf{x}^T P_i \mathbf{x} \geq 0$, we can derive that $\mathbf{q}_i^T\mathbf{x}+r_i \leq 0$, i.e., $\mathbf{q}_i^T\mathbf{x}+r_i-\frac{1}{2} \leq 0$.

Then (128) can be formulated as

$$\| A_i'\mathbf{x} + \mathbf{b}_i \|^2 \leq (\mathbf{q}_i^T\mathbf{x} + r_i - \frac{1}{2})^2 \tag{196}$$

$$\Leftrightarrow \| A_i'\mathbf{x} + \mathbf{b}_i \| \leq -(\mathbf{q}_i^T\mathbf{x} + r_i - \frac{1}{2}) \tag{197}$$