Optimization Theory and Methods

Prof. Xiaoe Ruan

Email: wruanxe@xjtu.edu.cn

Tel:13279321898

2019, Fall

Chapter 6. Linearly Constrained Optimization Methods

1. Projected Gradient Method and Reduced Gradient Method
2. Constrained Null Space Method
3. Active Set Method
4. Quadratic Programming

Nonlinear Optimization with Linear Constraints（NOLC）Model:

$$\begin{cases} \min\ f(x), \quad \text{s.t.} \\ a_i^{\mathrm{T}} x = b_i, i \in E = \{1,2,\cdots m_e\}, \\ a_i^{\mathrm{T}} x \geq b_i, i \in I = \{m_e + 1, \cdots m\}. \end{cases}$$

$f(x)$ **--Nonlinear smooth fcn.** $a_i \in R^n, b_i \in R, i = 1, \cdots, m.$

$$D = \left\{ x \middle| a_i^{\mathrm{T}} x = b_i, i \in E, a_i^{\mathrm{T}} x \geq b_i, i \in I \right\}$$

--Constraint Set, Constraint Domain or Feasible Domain.

Optimization Strategy:

Construct iterative sequence $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$. **s.t.**

(1) $f\left(x^{(k)} + \alpha_k d^{(k)}\right) < f\left(x^{(k)}\right)$. **(2)** $x^{(k)} + \alpha_k d^{(k)} \in D$.

1. Rosen Projected Gradient Algorithm

Given matrix $P = P^{\mathrm{T}}$ **and linear transform** $y = Px, \forall x \in R^n$.

If $y^{\mathrm{T}}(x - Px) = 0$, **then** $y = Px$ **is a projection transform.**

P **---projection matrix.** **Properties:**

$P = [P_1 \mid P_2 \mid \cdots \mid P_n]$

(1) $P^2 = P$. **(2)** P **is semi-PD.**

(3) $\Longleftrightarrow$ $I - P$ **is a projection matrix.**

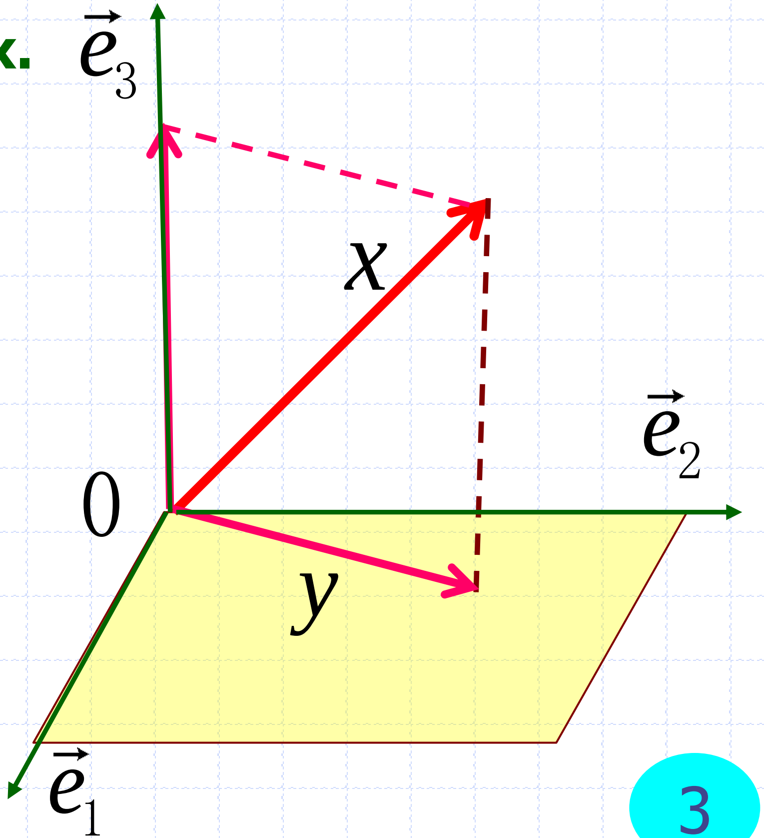
(4) **Subspace** $L = \{Px, x \in R^n\}$

$= \{P_1 x_1 + P_2 x_2 + \cdots + P_n x_n, x \in R^n\}$

is orthogonal to subspace

$L^{\perp} = \{x - Px, x \in R^n\}$. **e.g.**

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad Px = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}.$$



Lemma1 Suppose $x^{(k)}$ is a feasible point of NOLC model.

Denote $I_e^{(k)} = I_e\left(x^{(k)}\right) = \left\{i \middle| a_i^{\mathrm{T}} x^{(k)} = b_i, i \in E\right\} = E = \{1,2,\cdots m_e\}$,

$$I_I^{(k)} = I_I\left(x^{(k)}\right) = \left\{i + m_e \middle| a_{i+m_e}^{\mathrm{T}} x^{(k)} = b_{i+m_e}, i + m_e \in I\right\},$$

$$I^{(k)} = I_e^{(k)} \bigcup I_I^{(k)}, \qquad A_{I^{(k)}} = \left[a_1 | \cdots | a_{m_e} | a_{m_e+1} | \cdots | a_{m_k}\right]$$

If the column rank of $A_{I^{(k)}}$ is full, **then**

（1） $P = I - A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}$ is a projection matrix.

（2） If $P\nabla f\left(x^{(k)}\right) \neq 0$, then $d^{(k)} = -P\nabla f\left(x^{(k)}\right)$ is feasible descending direction of NOLC at $x^{(k)}$.

Proof:（1） $I - P = A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}$ **then**

$$(I-P)^2 = \left[A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}\right]\left[A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}\right] = I - P$$

(2) $$\nabla f\left(x^{(k)}\right)^{\mathrm{T}} d^{(k)} = -\nabla f\left(x^{(k)}\right)^{\mathrm{T}} P \nabla f\left(x^{(k)}\right) = -\nabla f\left(x^{(k)}\right)^{\mathrm{T}} P^{\mathrm{T}} P \nabla f\left(x^{(k)}\right)$$

$$= -\left\| P\nabla f\left(x^{(k)}\right)\right\|^2 < 0.$$ **i.e.** $d^{(k)} = -P\nabla f\left(x^{(k)}\right)$ **is descending direction.**

In addition $A_{I^{(k)}} = \left[a_1 \mid a_2 \mid \cdots \mid a_{m_k}\right]$, $P = I - A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}.$

Therefore $$A_{I^{(k)}}^{\mathrm{T}} d^{(k)} = -A_{I^{(k)}}^{\mathrm{T}}\left[I - A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}} A_{I^{(k)}}\right)^{-1} A_{I^{(k)}}^{\mathrm{T}}\right]\nabla f\left(x^{(k)}\right) = 0$$

Thus **a. If** $i \in I^{(k)}$, **then for any step** $\alpha_k > 0$,

$$a_i^{\mathrm{T}}\left(x^{(k)} + \alpha_k d^{(k)}\right) = b_i$$ **holds.**

b. If $i \in E \bigcup I \setminus I^{(k)}$, **i.e.** $a_i^{\mathrm{T}} x^{(k)} > b_i$. **Choose** $0 < \alpha_k << 1$,

s.t. $a_i^{\mathrm{T}}\left(x^{(k)} + \alpha_k d^{(k)}\right) > b_i$ **guaranteed.**

i.e. **If** $P\nabla f\left(x^{(k)}\right) \neq 0$,

then $d^{(k)} = -P\nabla f\left(x^{(k)}\right)$ **is feasible descending direction.**

Lemma2 **Assume that (A1)** $x^{(k)}$ **is a feasible point of NOLC model.**

$$I_e^{(k)}=I_e\left(x^{(k)}\right)=\left\{i \middle| a_i^{\mathrm{T}}x^{(k)}=b_i, i\in E\right\} = E=\{1,2,\cdots m_e\}.$$

$$I_I^{(k)}=I_I\left(x^{(k)}\right)=\left\{i+m_e \middle| a_{i+m_e}^{\mathrm{T}}x^{(k)}=b_{i+m_e}, i+m_e\in I\right\},\qquad I^{(k)}=I_e^{(k)}\cup I_I^{(k)}.$$

(A2) Column rank of $A_{I^{(k)}}=\left[a_1|\cdots|a_{m_e}|a_{m_e+1}|\cdots|a_{m_k}\right]$ **is full,**

$P=I-A_{I^{(k)}}\left(A_{I^{(k)}}^{\mathrm{T}}A_{I^{(k)}}\right)^{-1}A_{I^{(k)}}^{\mathrm{T}}$ **is corresponding projection matrix and** $P\nabla f\left(x^{(k)}\right)=0.$

Denote $u=\left(A_{I^{(k)}}^{\mathrm{T}}A_{I^{(k)}}\right)^{-1}A_{I^{(k)}}^{\mathrm{T}}\nabla f\left(x^{(k)}\right)=\left[\underbrace{u_1,\cdots,u_{m_e}}_{v^{\mathrm{T}}}\,\middle|\,\underbrace{u_{m_e+1},\cdots,u_{m_k}}_{w^{\mathrm{T}}}\right]^{\mathrm{T}}=\begin{bmatrix}v\\w\end{bmatrix}$

Then **(1) If** $w\ge 0$**, then** $x^{(k)}$ **is KT point of NOLC model.**

(2) If $w_j<0$, **then** $\bar{P}=I-\bar{A}_{I^{(k)}}\left(\bar{A}_{I^{(k)}}^{\mathrm{T}}\bar{A}_{I^{(k)}}\right)^{-1}\bar{A}_{I^{(k)}}^{\mathrm{T}}$ **is projection matrix .**

$\bar{d}^{(k)}=-\bar{P}\nabla f\left(x^{(k)}\right)$ **is feasible descending direction of NOLC model.**

Here $\bar{A}_{I^{(k)}}=\left[a_1|\cdots|a_{m_e}|\cdots|a_{m_e+j-1}|a_{m_e+j+1}|\cdots|a_{m_k}\right]$

Recall1 (1st-order Kuhn-Tucker Necessity Conditions)

For (NP):

$$\begin{cases} \min \ f(x), \\ \text{s. t.} \quad c_i(x)=0, i\in E=\{1,2,\cdots m'\}, \\ \qquad\quad c_i(x)\ge 0, i\in I=\{m'+1,\cdots m\}. \end{cases}$$

Given x^* **is local minimizer of NP,** $f(x)$ **and** $c_i(x)(i\in E\cup I)$ **are 1-st-order continuously differentiable at** x^*

If $SFD(x^*,d)=LFD(x^*,d)$, **then there exists** $\lambda^*=(\lambda_1^*,\lambda_2^*,\cdots,\lambda_m^*)^{\mathrm{T}}$

Kuhn-Tucker point

Kuhn-Tucker cond.

s.t.

$$\nabla f(\boxed{x^*})-\sum_{i=1}^{m}\lambda_i^*\nabla c_i(x^*)=0, i\in E\cup I,$$

$$\lambda_i^*\ge 0, i\in I, \quad \boxed{\lambda_i^* c_i(x^*)=0, i\in I.}$$

Complementary relaxed cond.

Proof:（1） From Assumptions (A1) and (A2),

For NP:

$$\begin{cases} \min\ f(x), \quad \text{s.t.} \\ a_i^{\mathrm{T}} x^{(k)} = b_i, i \in E = \{1,2,\cdots m_e\}, \\ a_i^{\mathrm{T}} x^{(k)} \ge b_i, i \in I = \{m_e+1,\cdots m\}. \end{cases}$$

$$I_e^{(k)} = I_e\left(x^{(k)}\right) = \left\{ i \middle| a_i^{\mathrm{T}} x^{(k)} = b_i, i \in E \right\} = E = \{1,2,\cdots m_e\},$$

$$I_I^{(k)} = I_I\left(x^{(k)}\right) = \left\{ i+m_e \middle| a_{i+m_e}^{\mathrm{T}} x^{(k)} = b_{i+m_e}, i+m_e \in I \right\}, \quad I^{(k)} = I_e^{(k)} \bigcup I_I^{(k)}.$$

We have $0 = P\nabla f\left(x^{(k)}\right) = \nabla f\left(x^{(k)}\right) - A_{I^{(k)}} u$

$$= \nabla f\left(x^{(k)}\right) - \sum_{i=1}^{m_e} v_i a_i - \sum_{j=m_e+1}^{m_k} w_{j-m_e} a_j - \sum_{j=m_k+1}^{m} 0 \cdot a_j.$$

and $w_{j-m_e}\left(a_j^{\mathrm{T}} x^{(k)} - b_j\right) = 0, \quad j = m_e+1,\cdots,m_k,$

$0 \cdot \left(a_j^{\mathrm{T}} x^{(k)} - b_j\right) = 0, \quad j = m_k+1,\cdots,m.$

i.e. $x^{(k)}$ **is KT pnt.**

Proof: (2) $\bar{A}_{I(k)} = \left[a_1 | \cdots | a_{m_e} | \cdots | a_{m_e+j-1} | a_{m_e+j+1} | \cdots | a_{m_k} \right],$

$$\bar{P} = I - \bar{A}_{I(k)} \left(\bar{A}_{I(k)}^{\mathrm{T}} \bar{A}_{I(k)} \right)^{-1} \bar{A}_{I(k)}^{\mathrm{T}},$$

Let $\bar{u} = \left(\bar{A}_{I(k)}^{\mathrm{T}} \bar{A}_{I(k)} \right)^{-1} \bar{A}_{I(k)}^{\mathrm{T}} \nabla f\left(x^{(k)}\right)$. **Contrarily suppose that**

$$0 = \bar{P} \nabla f\left(x^{(k)}\right) = \left[I - \bar{A}_{I(k)} \left(\bar{A}_{I(k)}^{\mathrm{T}} \bar{A}_{I(k)} \right)^{-1} \bar{A}_{I(k)}^{\mathrm{T}} \right] \nabla f\left(x^{(k)}\right)$$

$$= \nabla f\left(x^{(k)}\right) - \bar{A}_{I(k)} \bar{u} = \nabla f\left(x^{(k)}\right) - \sum_{i=1}^{m_e} \bar{u}_i a_i - \sum_{i=m_e+1, i \neq m_e+j}^{m_k} \bar{u}_i a_i .$$

Then $\displaystyle \sum_{i=1}^{m_e} (u_i - \bar{u}_i) a_i + \sum_{i=m_e+1, i \neq m_e+j}^{m_k} \left(w_{i-m_e} - \bar{u}_i \right) a_i + w_j a_{m_e+j} = 0.$

This contradicts to the full column rank of matrix $A_{I(k)}$.

Analogous to Lemma1, it is easy to prove $\bar{P}$ **is a projection matrix and** $\bar{d}^{(k)} = -\bar{P}\nabla f\left(x^{(k)}\right)$ **is descending direction.**

Feasibility: First, similar to Lemma1, it is easy to prove $\bar{A}_{I^{(k)}}^{\mathrm{T}}\bar{d}^{(k)}=0.$

Thus **a. If** $i \in I^{(k)} \setminus \{m_e + j\}$ **then for any** $\bar{\alpha}_k > 0,$

$$a_i^{\mathrm{T}}\left(x^{(k)} + \bar{\alpha}_k \bar{d}^{(k)}\right) = b_i$$ **holds.**

b. If $i \in E \bigcup I \setminus I^{(k)}$ **then** $a_i^{\mathrm{T}} x^{(k)} > b_i$. **Choose** $0 < \bar{\alpha}_k << 1,$

s.t. $a_i^{\mathrm{T}}\left(x^{(k)} + \bar{\alpha}_k \bar{d}^{(k)}\right) > b_i$ **is guaranteed.**

c. If $i = m_e + j,$ **recalling** $\nabla f\left(x^{(k)}\right) - A_{I^{(k)}} u=0$ **induces**

$$a_{m_e+j}^{\mathrm{T}}\bar{d}^{(k)} = -a_{m_e+j}^{\mathrm{T}}\bar{P}\nabla f\left(x^{(k)}\right) = -a_{m_e+j}^{\mathrm{T}}\bar{P}A_{I^{(k)}}u = -a_{m_e+j}^{\mathrm{T}}\bar{P}\left[\bar{A}_{I^{(k)}}\hat{u} + w_j a_{m_e+j}\right]$$

$$= -a_{m_e+j}^{\mathrm{T}}\left[I - \bar{A}_{I^{(k)}}\left(\bar{A}_{I^{(k)}}^{\mathrm{T}}\bar{A}_{I^{(k)}}\right)^{-1}\bar{A}_{I^{(k)}}^{\mathrm{T}}\right]\bar{A}_{I^{(k)}}\hat{u} - w_j a_{m_e+j}^{\mathrm{T}}\bar{P}a_{m_e+j}$$

$$= -w_j a_{m_e+j}^{\mathrm{T}}\bar{P}a_{m_e+j} \geq 0.$$ **Thus** $a_{m_e+j}^{\mathrm{T}}\left(x^{(k)} + \bar{\alpha}_k \bar{d}^{(k)}\right) \geq b_i \quad \left(\forall \bar{\alpha}_k > 0\right)$

$$d^{(k)} = -P\nabla f\left(x^{(k)}\right) \quad \text{or} \quad d^{(k)} = -\bar{P}\nabla f\left(x^{(k)}\right)$$

Determination of α: **Suppose that** $x^{(k)}$ **is a feasible but not KT pnt.**

i.e. $$\begin{cases} a_i^{\mathrm{T}} x^{(k)} = b_i, i \in E = \{1,2,\cdots m_e\}, \\ a_i^{\mathrm{T}} x^{(k)} \geq b_i, i \in I = \{m_e+1,\cdots m\} \end{cases}$$

From Lemma1 & 2

a. While $i \in I_e^{(k)} = \left\{ i \middle| a_i^{\mathrm{T}} x^{(k)} = b_i, i \in E \right\} = E = \{1,2,\cdots m_e\}$

for any step $\alpha > 0$, $a_i^{\mathrm{T}}\left(x^{(k)} + \alpha d^{(k)}\right) = b_i$ **holds.**

b. While $i \in E \bigcup I \setminus I_e^{(k)}$, **i.e.** $a_i^{\mathrm{T}} x^{(k)} \geq b_i$.

If $a_i^{\mathrm{T}} d^{(k)} \geq 0$, **then for any** $\alpha > 0$, $a_i^{\mathrm{T}}\left(x^{(k)} + \alpha d^{(k)}\right) \geq b_i$ **true.**

If $a_i^{\mathrm{T}} d^{(k)} < 0$ **and** $a_i^{\mathrm{T}}\left(x^{(k)} + \alpha d^{(k)}\right) \geq b_i$, **then** $0 < \alpha < \dfrac{a_i^{\mathrm{T}} x^{(k)} - b_i}{-a_i^{\mathrm{T}} d^{(k)}}$.

Hence $$\alpha_{max} = min\left\{ \frac{a_i^{\mathrm{T}} x^{(k)} - b_i}{-a_i^{\mathrm{T}} d^{(k)}} \middle| a_i^{\mathrm{T}} d^{(k)} < 0, i \in E \bigcup I \setminus I_e^{(k)} \neq \varnothing \right\}$$ **or** $\alpha_{max} = +\infty$.

In 1989, Du and Zhang proved that Rosen Projected Gradient Algorithm is globally convergent.

(2) Wolfe Reduced Gradient Algorithm

Define relaxed variable $z_i \geq 0, \quad i \in I = \{m_e + 1, \cdots m\}$.

Then $a_i^{\mathrm{T}} x \geq b_i, i \in I = \{m_e + 1, \cdots m\}$

$$\Leftrightarrow \quad a_i^{\mathrm{T}} x - z_i = b_i, i \in I = \{m_e + 1, \cdots m\}$$

NOLC:

$$\begin{cases} \min\ f(x), \quad \text{s. t.} \\ a_i^{\mathrm{T}} x = b_i, i \in E = \{1, 2, \cdots m_e\}, \\ a_i^{\mathrm{T}} x \geq b_i, i \in I = \{m_e + 1, \cdots m\}. \end{cases}$$

$\Leftrightarrow$

NOLEC:

$$\begin{cases} \min\ f(x), \\ \text{s. t. } Ax = b, \\ \qquad x \geq 0. \end{cases}$$

Def.1: Basis(Basis Matrix) ——If a square submatrix B of A is invertible, then matrix B is said as a basis (basis matrix).

Given

$$A=\begin{array}{cccccc} P_1 & \cdots & P_m & P_{m+1} & \cdots & P_n \end{array}$$

$$A=\left[\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1m} & a_{1,m+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2m} & a_{2,m+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mm} & a_{m,m+1} & \cdots & a_{mn} \end{array}\right],$$

$$B=\left[P_1\,|\cdots|\,P_m\right],$$

$$N=\left[P_{m+1}\,|\cdots|\,P_n\right],$$

Basis vectors: $P_1,\cdots,P_m$. **Non-basis vectors:** $P_{m+1},\cdots,P_n$.

Basis variable: $x_B=\left[x_1,\cdots,x_m\right]^{\mathrm{T}}$,

Non-basis variable: $x_N=\left[x_{m+1},\cdots x_n\right]^{\mathrm{T}}$,

$$x=\left[x_B^{\mathrm{T}}\,|\,x_N^{\mathrm{T}}\right]^{\mathrm{T}}.$$

Then $A=[B\,|\,N], x=\begin{bmatrix} x_B \\ x_N \end{bmatrix}$. **And** $Ax=b \Rightarrow Bx_B+Nx_N=b$

$$\Rightarrow \quad x_B(x_N)=B^{-1}b-B^{-1}Nx_N$$

$$f(x)=f\left(x_B(x_N),x_N\right) \triangleq F(x_N)$$ **Thus**

NOLEC: $\begin{cases} \min\ f(x), \\ \text{s. t. } Ax=b, \\ \qquad x\geq 0. \end{cases} \Longleftrightarrow$ **NONNC:** $\begin{cases} \min\ F(x_N), \\ \text{s. t. } x_B(x_N)\geq 0, \\ \qquad x_N\geq 0. \end{cases}$

$$r(x_N) \triangleq \nabla F(x_N) = \nabla f\left(x_B(x_N),x_N\right)$$

$$=\nabla_{x_N} f\left(x_B(x_N),x_N\right) - \left(B^{-1}N\right)^{\mathrm{T}} \nabla_{x_B} f\left(x_B(x_N),x_N\right)$$

--Reduced gradient

Lemma3 Suppose that $x^{(k)} = \left[x_B^{(k)\mathrm{T}} \mid x_N^{(k)\mathrm{T}} \right]^{\mathrm{T}} \quad \left(x_B^{(k)} > 0 \right)$

is a nonzero feasible point of NOLEC $\min f(x), \text{s.t. } Ax = b,\ x \geq 0.$

Its reduced gradient $r\left(x_N^{(k)}\right) = \nabla F\left(x_N^{(k)}\right) = \left[r_1\left(x_N^{(k)}\right), \cdots, r_{n-m}\left(x_N^{(k)}\right) \right]^{\mathrm{T}}$

Let $d^{(k)} = \left[d_B^{(k)\mathrm{T}} \mid d_N^{(k)\mathrm{T}} \right]^{\mathrm{T}},$ where $d_B^{(k)} = -B^{-1} N d_N^{(k)},$

$$d_N^{(k)} = \left[d_{N_1}^{(k)}, \cdots, d_{N_{n-m}}^{(k)} \right]^{\mathrm{T}}$$ with $$d_{N_j}^{(k)} = \begin{cases} -x_{N_j}^{(k)} r_j\left(x_N^{(k)}\right), & if \quad r_j\left(x_N^{(k)}\right) > 0, \\ -r_j\left(x_N^{(k)}\right), & if \quad r_j\left(x_N^{(k)}\right) \leq 0; \end{cases}$$

Then

(1) If $d^{(k)} \neq 0, \Longrightarrow d^{(k)}$ **is a feasible descending direction of NOLEC.**

(2) If $d^{(k)} = 0 \Longleftrightarrow x^{(k)}$ **is KT point of NOLEC.**

Proof:（1） $Ad^{(k)} = Bd_B^{(k)} + Nd_N^{(k)} = B\left(-B^{-1}Nd_N^{(k)}\right) + Nd_N^{(k)} = 0$

Then $Ax^{(k+1)} = A\left(x^{(k)} + \alpha d^{(k)}\right) = b.$ **For non-basis vector** $x_N^{(k)}$,

If $r_{N_j}\left(x_N^{(k)}\right) \le 0,$ **then for any** $\alpha > 0,$ **we have** $x_{N_j}^{(k)} + \alpha d_{N_j}^{(k)} \ge 0.$

If $r_{N_j}\left(x_N^{(k)}\right) > 0,\ x_{N_j}^{(k)} > 0,$ **properly choose** $\alpha > 0,$ **s.t.** $x_{N_j}^{(k)} + \alpha d_{N_j}^{(k)} \ge 0.$

If $r_{N_j}\left(x_N^{(k)}\right) > 0,\ x_{N_j}^{(k)} = 0,$ **then** $d_{N_j}^{(k)} = 0.$ **i.e.** $x_{N_j}^{(k)} + \alpha d_{N_j}^{(k)} = 0.$

Therefore $x_B^{(k)} > 0,$ **Properly choose** $\alpha > 0,$ **s.t.** $x_B^{(k)} + \alpha d_B^{(k)} \ge 0.$

This means that $d^{(k)}$ **is feasible direction.**

Besides $\nabla f\left(x^{(k)}\right)^{\mathrm{T}} d^{(k)} = \nabla_{x_B^{(k)}} f\left(x^{(k)}\right)^{\mathrm{T}} d_B^{(k)} + \nabla_{x_N^{(k)}} f\left(x^{(k)}\right)^{\mathrm{T}} d_N^{(k)}$

$$= \left[-\nabla_{x_B^{(k)}} f\left(x^{(k)}\right)^{\mathrm{T}} B^{-1}N + \nabla_{x_N^{(k)}} f\left(x^{(k)}\right)^{\mathrm{T}}\right] d_N^{(k)} = r\left(x_N^{(k)}\right)^{\mathrm{T}} d_N^{(k)} < 0.$$

This implies that $d^{(k)}$ **is descending direction.**

Proof : (2) $d^{(k)} = 0 \Longleftrightarrow x^{(k)}$ **is KT pnt of NOLEC**

$\Longleftarrow$

Let $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$, **then** $\begin{cases} \min\ f(x), \\ \text{s. t. } Ax = b, \\ \qquad x \geq 0. \end{cases} \Longleftrightarrow \begin{cases} \min\ f(x), \text{s. t.} \\ b_j - A_j x = 0, j = 1, \cdots, m; \\ x_i \geq 0, i = 1, \cdots, n \end{cases}$

Construct Lagrange fcn: $L(x,v,u) = f(x) - \sum_{j=1}^{m} v_j (b_j - A_j x) - \sum_{i=1}^{n} u_i x_i$

Then $\nabla_x L(x^{(k)}, v, u) = \begin{bmatrix} \nabla_{x_B} f(x^{(k)}) \\ \nabla_{x_N} f(x^{(k)}) \end{bmatrix} + \begin{bmatrix} B^{\mathrm{T}} \\ N^{\mathrm{T}} \end{bmatrix} v - \begin{bmatrix} u_B \\ u_N \end{bmatrix} = 0, \ u = \begin{bmatrix} u_B \\ u_N \end{bmatrix} \geq 0,$

and $u_{B_i} x_{B_i}^{(k)} = 0, \ u_{N_j} x_{N_j}^{(k)} = 0.$ **As** $x_B^{(k)} > 0$ **implies** $u_B = 0.$

Thus $\nabla_{x_B} f(x^{(k)}) + B^{\mathrm{T}} v = 0$ **i.e.** $v = -(B^{\mathrm{T}})^{-1} \nabla_{x_B} f(x^{(k)})$ **Hence**

$u_N = \nabla_{x_N} f(x^{(k)}) - (B^{-1} N)^{\mathrm{T}} \nabla_{x_B} f(x^{(k)}) = r(x_N^{(k)}) \geq 0.$ **But** $u_{N_j} x_{N_j}^{(k)} = 0,$

Therefore $x_j^{(k)} = 0$ **or** $u_{N_j} = r_{N_j}(x^{(k)}) = 0.$ **i.e.** $d^{(k)} = 0.$

Mind $d_{N_j}^{(k)} = \begin{cases} -x_{N_j}^{(k)} r_j\left(x_N^{(k)}\right), & if \quad r_j\left(x_N^{(k)}\right) > 0, \\ -r_j\left(x_N^{(k)}\right), & if \quad r_j\left(x_N^{(k)}\right) \le 0; \end{cases}$ $d^{(k)} = \begin{bmatrix} d_B^{(k)} \\ d_N^{(k)} \end{bmatrix}.$

$$d_N^{(k)} = \left[d_{N_1}^{(k)}, \cdots, d_{N_{n-m}}^{(k)}\right]^{\mathrm{T}}, \quad d_B^{(k)} = -B^{-1} N d_N^{(k)},$$

If $d^{(k)} = 0$ **then** $r_j\left(x_N^{(k)}\right) = 0$ **or** $x_{N_j}^{(k)} r_j\left(x_N^{(k)}\right) = 0 \left(if \quad r_j\left(x_N^{(k)}\right) > 0\right)$

If $r_j\left(x_N^{(k)}\right) = 0$ **Choose** $u_{N_j} = 0$

If $r_j\left(x_N^{(k)}\right) > 0$ **then** $x_{N_j}^{(k)} = 0$ **Choose** $u_{N_j} = r_j\left(x_N^{(k)}\right) > 0$

Let $u_B = 0, \quad v = -\left(B^{\mathrm{T}}\right)^{-1} \nabla_{x_B} f\left(x^{(k)}\right)$ **Then**

$$\nabla_x L\left(x^{(k)}, v, u\right) = \begin{bmatrix} \nabla_{x_B} f\left(x^{(k)}\right) \\ \nabla_{x_N} f\left(x^{(k)}\right) \end{bmatrix} + \begin{bmatrix} B^{\mathrm{T}} \\ N^{\mathrm{T}} \end{bmatrix} v - \begin{bmatrix} u_B \\ u_N \end{bmatrix} = 0, \quad u = \begin{bmatrix} u_B \\ u_N \end{bmatrix} \ge 0,$$

$u_i x_i^{(k)} = 0, i = 1, \cdots, n.$ **holds.** **i.e.** $x^{(k)}$ **is KT point.**

Above discussion results in

NOLEC: $\begin{cases} \min\ f(x), \\ \text{s. t. } Ax = b, \\ \qquad x \geq 0. \end{cases}$ $\Longleftrightarrow$ NONNC: $\begin{cases} \min\ F(x_N), \\ \text{s. t. } x_B(x_N) \geq 0, \\ \qquad x_N \geq 0. \end{cases}$

Then, the conclusions

(1) If $d^{(k)} \neq 0,\quad d^{(k)}$ **is descending direction of NOLEC.**

(2) $d^{(k)} = 0$ **iff** $x^{(k)}$ **is KT point of NOLEC.**

$\Longleftrightarrow$

(1) If $d_N^{(k)} \neq 0,\ d_N^{(k)}$ **is descending direction of NONNC.**

(2) $d_N^{(k)} = 0$ **iff** $x_N^{(k)}$ **is KT point of NONNC.**

KT conditions of NONNC: $r\left(x_N^{(k)}\right) \geq 0$ **and** $r\left(x_N^{(k)}\right)^{\mathrm{T}} x_N^{(k)} = 0.$

Determination of step α: $\quad x^{(k)} + \alpha d^{(k)} \geq 0.$

$$\alpha_{max} = min\left\{ -\frac{x_j^{(k)}}{d_j^{(k)}} \middle| d_j^{(k)} < 0, \{j\} \neq \varnothing \right\}$$ **or** $$\alpha_{max} = +\infty.$$

For NOLEC: $\quad \mathtt{min}\ f(x), \mathtt{s.\ t.}\ Ax = b, x \geq 0.$

Convergence conclusion

If any *m* columns of matrix *A* are independent, then any accumulation point of the sequence, which is generated by the reduced gradient algorithm starting from a nonzero basis feasible point, is KT point.

Generalization:

$$\begin{cases} \mathtt{min}\ f(x), \\ \mathtt{s.\ t.}\ Ax = b, \\ \qquad x \geq 0. \end{cases} \Longrightarrow \begin{cases} \mathtt{min}\ \nabla f\left(x^{(k)}\right)^{\mathrm{T}} x, \\ \mathtt{s.\ t.}\ Ax = b, \\ \qquad x \geq 0. \end{cases}$$

2. Description of Null Space

Suppose that column rank of matrix $A=\left[a_1\,|\,a_2\,|\cdots|\,a_m\right]$ **is full.**

Then

NOLEC $\begin{cases}\min\ f(x), \quad \text{s. t.} \\ a_i^{\mathrm{T}}x=b_i, i=1,2,\cdots m.\end{cases} \Rightarrow \begin{cases}\min\ f(x), \\ \text{s. t. } A^{\mathrm{T}}x=b.\end{cases}$

Let $V=\left\{k_1a_1+k_2a_2+\cdots+k_ma_m \,\middle|\, k_i\in R\right\}$. **Basis:** $S_1, S_2, \cdots, S_m$.

Denote $W=\left\{y \,\middle|\, A^{\mathrm{T}}y=0 \,\middle|\, y\in R^n\right\}$ **--Constrained Null Space.**

$\dim W = n-m$ **and** $W\perp V$, **Basis:** $Z_1, Z_2, \cdots, Z_{n-m}$.

Let $S=\left[S_1\,|\,S_2\,|\cdots|\,S_m\right]$, $Z=\left[Z_1\,|\,Z_2\,|\cdots|\,Z_{n-m}\right]$. **Then** $S^{\mathrm{T}}Z=0$.

$\left\{S_1, S_2, \cdots, S_m\right\}$ **is equivalent to** $\left\{a_1, a_2, \cdots, a_m\right\}$

Thus $A^{\mathrm{T}}Z=0$, **and** $A^{\mathrm{T}}S$ **nonsingular. Simply set** $A^{\mathrm{T}}S=I$.

As $S_1, S_2, \cdots, S_m, Z_1, Z_2, \cdots, Z_{n-m}$ **is a set of basis of** R^n

If $x = \hat{y}_1 S_1 + \hat{y}_2 S_2 + \cdots + \hat{y}_m S_m + \hat{x}_1 Z_1 + \hat{x}_2 Z_2 + \cdots + \hat{x}_{n-m} Z_{n-m}$

$= [S \mid Z] \begin{bmatrix} \hat{y} \\ \hat{x} \end{bmatrix} = S\hat{y} + Z\hat{x}$ **is a feasible point, then**

$A^{\mathrm{T}} x = A^{\mathrm{T}} (S\hat{y} + Z\hat{x}) = A^{\mathrm{T}} S\hat{y} + A^{\mathrm{T}} Z\hat{x} = \hat{y} = b$ **and**

$x = S\hat{y} + Z\hat{x} = Sb + Z\hat{x}.$ **Thus** $f(x) = f(Sb + Z\hat{x}) \triangleq F(\hat{x})$

NOLEC: $\begin{cases} \min\ f(x), \quad \text{s.t.} \\ a_i^{\mathrm{T}} x = b_i, i = 1, 2, \cdots m. \end{cases} \Rightarrow \min\limits_{\hat{x} \in R^{n-m}} F(\hat{x}).$ **No constraint!**

Difference of basis $Z = [Z_1 \mid Z_2 \mid \cdots \mid Z_{n-m}]$ **produces different algorithms. May adopt non-constraint optimization algorithm.**

Method1: $$A = Q\begin{bmatrix} R_{m\times m} \\ 0 \end{bmatrix} = [Q_1 \mid Q_2]\begin{bmatrix} R_{m\times m} \\ 0 \end{bmatrix} = Q_1 R.$$

- Q is orthogonal. i.e. $Q^{\mathrm{T}}Q = I_{n\times n}$, $Q_1^{\mathrm{T}}Q_1 = I_{m\times m}$, $Q_1^{\mathrm{T}}Q_2 = 0.$

  $R_{m\times m}$ is upper triangular and invertible.

Let $\boxed{S = Q_1 R^{-\mathrm{T}}}$, $\boxed{Z = Q_2}$. **Then**

$$A^{\mathrm{T}}Z = R^{\mathrm{T}}Q_1^{\mathrm{T}}Q_2 = 0, \quad A^{\mathrm{T}}S = R^{\mathrm{T}}Q_1^{\mathrm{T}}Q_1R^{-\mathrm{T}} = I.$$

$$A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}} = Q_1R\left(R^{\mathrm{T}}Q_1^{\mathrm{T}}Q_1R\right)^{-1}R^{\mathrm{T}}Q_1^{\mathrm{T}} = Q_1Q_1^{\mathrm{T}},$$

$$QQ^{\mathrm{T}} = [Q_1 \mid Q_2]\begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{bmatrix} = Q_1Q_1^{\mathrm{T}} + Q_2Q_2^{\mathrm{T}} = I$$ **and**

$$ZZ^{\mathrm{T}} = Q_2Q_2^{\mathrm{T}} = I - Q_1Q_1^{\mathrm{T}} = I - A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}.$$

Projection of ZZ^{T} **is obvious and thus Rosen PG adoptable.**

3. Active Set Method

Find the searching direction by reforming NOLC to NOLEC or only considering equality constraints at the point $x^{(k)}$.

i.e.

NOLC $\begin{cases} \min \ f(x), \quad \text{s. t.} \\ a_i^{\mathrm{T}} x = b_i, i \in E = \{1,2,\cdots m_e\}, \\ a_i^{\mathrm{T}} x \geq b_i, i \in I = \{m_e+1,\cdots m\}. \end{cases}$ $\Rightarrow$ **NOLEC** $\begin{cases} \min \ f(x), \quad \text{s. t.} \\ a_i^{\mathrm{T}} x = b_i, i = 1,2,\cdots m. \end{cases}$

Idea of ASM: **Given** $x^{(k)}$ **be a feasible pnt of NOLEC.**

Solving $\left\{ \min \ f\left(x^{(k)} + d\right), \text{s. t. } a_i^{\mathrm{T}} d = 0, i = 1,2,\cdots m. \right.$ **yields** $d^{(k)}$

If $d^{(k)} = 0$ **then** $x^{(k)}$ **is the minimizer of NOLEC.**

If $d^{(k)} \neq 0$ **then** $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$. **Find** $d^{(k+1)}$

Given the feasible point $x^{(k)}$, achieve $x^{(k+1)}$ following

Solve NOLEC:

$$\begin{cases} \min \frac{1}{2} x^{\mathrm{T}} G x + p^{\mathrm{T}} x & (2.1) \\ s.t. \quad a_i^{\mathrm{T}} x = b_i, \quad i \in E \cup I(x^{(k)}) = S_k & (2.2) \end{cases}$$

Let $\delta = x - x^{(k)}$ (2.1)+(2.2) becomes

$$\begin{cases} \min \frac{1}{2} \delta^{\mathrm{T}} G \delta + p_k^{\mathrm{T}} \delta & (2.3) \\ s.t. \quad a_i^{\mathrm{T}} \delta = 0, \quad i \in S_k & (2.4) \end{cases}$$

Here $p_k = Gx^{(k)} + p$

Solving(2.3)+(2.4) yields $\delta^{(k)}$

Case1: $\delta^{(k)} \neq 0$

If $x^{(k)} + \delta^{(k)}$ **is a feasible point.** **Let** $x^{(k+1)} = x^{(k)} + \delta^{(k)}$

If $x^{(k)} + \delta^{(k)}$ **is not a feasible point.** **Then for an index** $j \in I \setminus S_k$

s.t. $a_j^{\mathrm{T}}(x^{(k)} + \delta^{(k)}) < b_j$ **and** $a_j^{\mathrm{T}}\delta^{(k)} < 0$

Choose α_k **s.t.** $x^{(k+1)} = x^{(k)} + \alpha_k \delta^{(k)}$ **is a feasible point.**

i.e. $a_j^{\mathrm{T}}(x^{(k)} + \alpha_k \delta^{(k)}) \geq b_j$ **Denote** $\alpha_k = \min\limits_{\substack{j \in I \setminus S_k \\ a_j^{\mathrm{T}}\delta^{(k)} < 0}} \dfrac{b_j - a_j^{\mathrm{T}} x^{(k)}}{a_j^{\mathrm{T}} \delta^{(k)}}$

Then

$$\alpha_k = \min\left\{ \min_{\substack{j \in I \setminus S_k \\ a_j^{\mathrm{T}}\delta^{(k)} < 0}} \frac{b_j - a_j^{\mathrm{T}} x^{(k)}}{a_j^{\mathrm{T}} \delta^{(k)}}, 1 \right\} \qquad \textbf{(2.5)}$$

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

If $\alpha_k = \min\limits_{\substack{j \in I \setminus S_k \\ a_j^{\mathrm{T}} \delta^{(k)} < 0}} \dfrac{b_j - a_j^{\mathrm{T}} x^{(k)}}{a_j^{\mathrm{T}} \delta^{(k)}} = \dfrac{b_p - a_p^{\mathrm{T}} x^{(k)}}{a_p^{\mathrm{T}} \delta^{(k)}} < 1,$

$$a_p^{\mathrm{T}} x^{(k+1)} = a_p^{\mathrm{T}} x^{(k)} + \alpha_k a_p^{\mathrm{T}} \delta^{(k)} = b_p$$

Update $S_{k+1} = S_k \bigcup \{p\}$

If $\alpha_k = 1, \;\; S_{k+1} = S_k$

Case2: $\delta^{(k)} = 0$ Testify $x^{(k)}$ is the minimizer

Computing Lagrange multiplier: $\lambda_q^{(k)} = \min_{i \in I(x^{(k)}) \cap S_k} \lambda_i^{(k)}$

If $\lambda_q^{(k)} \geq 0$, from 1-st-order necessity $x^{(k)}$ is KT point.

If $\lambda_q^{(k)} < 0$, when $\tilde{x}$ deviates from line $a_q^{\mathrm{T}} x = b_q$ s.t.

$$a_q^{\mathrm{T}} \tilde{x} > b_q \quad \text{and} \quad a_i^{\mathrm{T}} \tilde{x} = b_i, \quad i \in S_k, i \neq q.$$

Let $d = \tilde{x} - x^{(k)}$ then $a_q^{\mathrm{T}} d > 0$ and $a_i^{\mathrm{T}} d = 0, \ i \in S_k, i \neq q.$

From $\nabla f(x^{(k)}) = \sum_{i \in S_k} \lambda_i^{(k)} a_i$ yields $d^{\mathrm{T}} \nabla f(x^{(k)}) = \sum_{i \in S_k} \lambda_i^{(k)} d^{\mathrm{T}} a_i = \lambda_q^{(k)} d^{\mathrm{T}} a_q < 0$

i.e. $d = \tilde{x} - x^{(k)}$ is a descending direction.

This means that the constraint $a_q^{\mathrm{T}} x = b_q$ is redundant.

Let $S_{k+1} = S_k \setminus \{q\}$, $x_{k+1} = x_k$ Find $\delta^{(k)}$.

Ex1: Find the Minimizer of following QP by ASM.

$$\min x_1^2 + x_2^2 - x_1 x_2 - 3x_1$$

$$\begin{aligned} s.t. \quad -x_1 - x_2 &\ge -2, \qquad (1) \\ x_1 &\ge 0, \qquad (2) \\ x_2 &\ge 0, \qquad (3) \end{aligned}$$

$$x^{(0)} = [0,0]^{\mathrm{T}}$$

Solution: **Objective fcn:** $f(x) = \dfrac{1}{2} x^{\mathrm{T}} G x + p^{\mathrm{T}} x$

where $G = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad p = \begin{bmatrix} -3 \\ 0 \end{bmatrix}.$

At point $x^{(0)} = [0,0]^{\mathrm{T}}$, **Active Set** $I_0 = \{2,3\}$

Let $p_0 = Gx^{(0)} + p, \quad S_0 = I_0 = \{2,3\}$

1-st iteration:

Solving NOLEC
$$\begin{cases} \min \varphi_0(\delta) = \frac{1}{2}\delta^{\mathrm{T}} G\delta + p_0^{\mathrm{T}}\delta, \\ s.t. \quad \delta_1 = 0, \quad \delta_2 = 0. \end{cases}$$
yields $\delta^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$

Compute Lagrange multiplier:

Let $L_0(\delta) = \frac{1}{2}\delta^{\mathrm{T}} G\delta + p_0^{\mathrm{T}}\delta - \lambda_2\delta_1 - \lambda_3\delta_2$

From 1-st-orderKuhn-Tucker necessity condition conducts

$$\begin{bmatrix} -3 \\ 0 \end{bmatrix} = \lambda_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
then $\lambda_2^{(0)} = -3, \quad \lambda_3^{(0)} = 0.$

Therefore $S_1 = S_0 \setminus \{2\} = \{3\}, \qquad x^{(1)} = x^{(0)} + \delta^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$

$$p_1 = Gx^{(1)} + p = \begin{bmatrix} -3 \\ 0 \end{bmatrix}.$$

2-nd iteration:

Solving NOLEC $\min \varphi_1(\delta) = \frac{1}{2}\delta^{T} G \delta + p_1^{T}\delta, \quad s.t. \quad \delta_2 = 0.$

makes $\delta^{(1)} = \begin{bmatrix} \frac{3}{2} \\ 0 \end{bmatrix} \neq 0$ **and** $x^{(1)} + \delta^{(1)} = \begin{bmatrix} \frac{3}{2} \\ 0 \end{bmatrix}$ **is feasible point !**

Let $x^{(2)} = x^{(1)} + \delta^{(1)} = \begin{bmatrix} \frac{3}{2} \\ 0 \end{bmatrix}$, $p_2 = Gx^{(2)} + p = \begin{bmatrix} 0 \\ -\frac{3}{2} \end{bmatrix}$,

$S_2 = S_1 = \{3\}$.

3-rd iteration:

Solving NOLEC $\min \varphi_2(\delta) = \frac{1}{2}\delta^{\mathrm{T}} G\delta + p_2^{\mathrm{T}}\delta, s.t. \quad \delta_2 = 0.$

obtains $\delta^{(2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Lagrange multiplier:

From $\begin{bmatrix} 0 \\ -\frac{3}{2} \end{bmatrix} = \lambda_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ **gives rise to** $\lambda_3^{(2)} = -\frac{3}{2}$

Therefore $S_3 = S_2 \setminus \{3\} = \Phi, \quad x^{(3)} = x^{(2)} = \begin{bmatrix} \frac{3}{2} \\ 0 \end{bmatrix}, \quad p_3 = p_2 = \begin{bmatrix} 0 \\ -\frac{3}{2} \end{bmatrix}$

4-th iteration:

Solving NO: $$\min \frac{1}{2}\delta^{\mathrm{T}} G\delta + p_3^{\mathrm{T}}\delta$$ **gets** $$\delta^{(3)} = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

$$x^{(3)} + \delta^{(3)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$ **is not feasible point**

$$\alpha_3 = \frac{b_1 - a_1^{\mathrm{T}} x^{(3)}}{a_1^{\mathrm{T}} \delta^{(3)}} = \frac{-2 + \frac{3}{2}}{-\frac{3}{2}} = \frac{1}{3}$$

Therefore $S_4 = S_3 \cup \{1\} = \{1\}$

$$x^{(4)} = x^{(3)} + \alpha_3 \delta^{(3)} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \end{bmatrix}, \qquad p_4 = Gx^{(4)} + p = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

5-th iteration:

Solving
$$\begin{cases} \min \dfrac{1}{2}\delta^{\mathrm{T}} G\delta + p_4^{\mathrm{T}}\delta, \\ \text{s.t.} \quad -\delta_1 - \delta_2 = 0. \end{cases}$$
captures
$$\delta^{(4)} = \begin{bmatrix} -\dfrac{1}{6} \\ \dfrac{1}{6} \end{bmatrix}$$

$$x^{(4)} + \delta^{(4)} = \begin{bmatrix} \dfrac{3}{2} \\ \dfrac{1}{2} \end{bmatrix}$$
is a feasible point. **Then**

$$x^{(5)} = x^{(4)} + \delta^{(4)} = \begin{bmatrix} \dfrac{3}{2} \\ \dfrac{1}{2} \end{bmatrix}, \quad S_5 = S_4 = \{1\}, \quad p_5 = Gx^{(5)} + p = \begin{bmatrix} -\dfrac{1}{2} \\ -\dfrac{1}{2} \end{bmatrix}$$

6-th iteration:

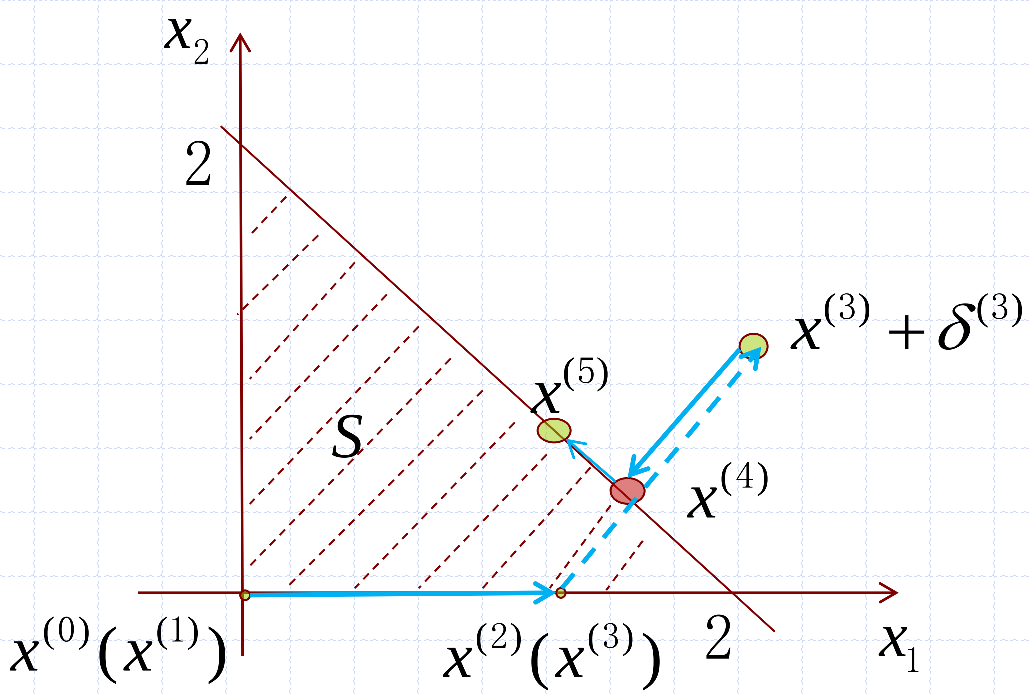
Solving
$$\min \frac{1}{2}\delta^{\mathrm{T}} G\delta + p_5^{\mathrm{T}}\delta$$
$$s.t. \quad -\delta_1 - \delta_2 = 0$$
makes $\delta^{(5)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Lagrange multiplier:

From $-\frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ **reaches** $\lambda_1^{(5)} = \frac{1}{2} > 0$

Therefore $x^{(5)} = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \end{bmatrix}$

is the minimizer.



4.Quadratic Programming

QPLC:

$$\begin{cases} \min\ q(x) = \dfrac{1}{2} x^{\mathrm{T}} G x + g^{\mathrm{T}} x, \\ \text{s. t.}\quad a_i^{\mathrm{T}} x = b_i, i \in E = \{1, 2, \cdots m_e\}, \\ \qquad\quad a_i^{\mathrm{T}} x \ge b_i, i \in I = \{m_e + 1, \cdots m\}. \end{cases}$$

where $G \in R^{n \times n}$ **is real symmetric.** $g \in R^n$

Find searching direction by introducing relaxing variable or only considering active inequality constraint.

i.e. QPLC is reformed to QPLEC.

QPLEC: $A = \left[a_1 \mid a_2 \mid \cdots \mid a_m\right]$ **full column rank**

QPLEC: $\min\ q(x) = \frac{1}{2}x^{\mathrm{T}}Gx + g^{\mathrm{T}}x,\ \text{s.t.}\ A^{\mathrm{T}}x = b.$

Elimination Method1

Let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix},\quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\quad G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix},\quad g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}.$

$A_1 \in R^{m\times m}$ **invertible** $\quad x_1, g_1 \in R^m,\quad x_2, g_2 \in R^{n-m},\quad G_{11} \in R^{m\times m},$

$G_{12} \in R^{m\times(n-m)},\quad G_{21} \in R^{(n-m)\times m},\quad G_{22} \in R^{(n-m)\times(n-m)},$

Then, from $A^{\mathrm{T}}x = \left[A_1^{\mathrm{T}} \mid A_2^{\mathrm{T}}\right]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A_1^{\mathrm{T}}x_1 + A_2^{\mathrm{T}}x_2 = b$

makes $x_1 = A_1^{-\mathrm{T}}\left(b - A_2^{\mathrm{T}}x_2\right)$

QPLEC: $\min\ q(x)=\frac{1}{2}x^{\mathrm{T}}Gx+g^{\mathrm{T}}x,\ \text{s.t.}\ A^{\mathrm{T}}x=b.$

$\Longrightarrow$ **QP1:** $\min_{x_2\in R^{n-m}}\ \hat{q}(x_2)=\frac{1}{2}x_2^{\mathrm{T}}\hat{G}x_2+\hat{g}^{\mathrm{T}}x_2+\hat{c}$

$$\hat{G}=G_{22}-G_{21}A_1^{-\mathrm{T}}A_2^{\mathrm{T}}-A_2A_1^{-1}G_{21}+A_2A_1^{-1}G_{11}A_1^{-\mathrm{T}}A_2^{\mathrm{T}},$$

$$\hat{g}=g_2-A_2A_1^{-1}g_1+\left(G_{21}-A_2A_1^{-1}G_{11}\right)A_1^{-1}b,\qquad \hat{c}=\frac{1}{2}b^{\mathrm{T}}A_1^{-1}G_{11}A_1^{-\mathrm{T}}b+g_1^{\mathrm{T}}A_1^{-\mathrm{T}}b.$$

(1) If $\hat{G}$ PD, the minimizer of QP1: $x_2^*=-\hat{G}^{-\mathrm{T}}\hat{g}$

Thus, the minimizer of QPLEC: $x^*=\begin{bmatrix}x_1^*\\ x_2^*\end{bmatrix}=\begin{bmatrix}A_1^{-\mathrm{T}}b+A_1^{-\mathrm{T}}A_2^{\mathrm{T}}\hat{G}^{-\mathrm{T}}\hat{g}\\ -\hat{G}^{-\mathrm{T}}\hat{g}\end{bmatrix}$

(2) If $\hat{G}$ exists negative eigenvalue,

then QP1 is not lower bounded.

i.e. QPLEC has no finite minimizer.

(3) If $\hat{G}$ **semi-PD.** **If** $\left(I-\hat{G}\hat{G}^{+}\right)\hat{g}=0$

the minimizer of QP1 $x_2^{*}=-\hat{G}^{+}\hat{g}+\left(I-\hat{G}^{+}\hat{G}\right)\tilde{x}$

Here $\hat{G}^{+}$ **is generalized inversion of** $\hat{G}$.

Then

$$x^{*}=\begin{bmatrix} x_1^{*} \\ x_2^{*} \end{bmatrix}=\begin{bmatrix} A_1^{-\mathrm{T}}b+A_1^{-\mathrm{T}}A_2^{\mathrm{T}}\hat{G}^{+}\hat{g}-A_1^{-\mathrm{T}}A_2^{\mathrm{T}}\left(I-\hat{G}^{+}\hat{G}\right)\tilde{x} \\ -\hat{G}^{+}\hat{g}+\left(I-\hat{G}^{+}\hat{G}\right)\tilde{x} \end{bmatrix}$$

If $\left(I-\hat{G}\hat{G}_{+}\right)\hat{g}\neq 0$ **QPLEC has no finite minimizer.**

Advantages: **Simple and explicit solution**

Disadvantages:

Unstable by larger computing error of matrix inversion.

Elimination Method 2

Adopt Null Space transform $x = S\hat{y} + Z\hat{x} = Sb + Z\hat{x}$

QPLEC: $\min\ q(x) = \frac{1}{2}x^{\mathrm{T}}Gx + g^{\mathrm{T}}x,\ \text{s.t.}\ A^{\mathrm{T}}x = b.$

$\Longrightarrow$ **QP2:** $\min_{\hat{x}\in R^{n-m}} \hat{q}(\hat{x}) = \frac{1}{2}\hat{x}^{\mathrm{T}}\hat{G}\hat{x} + \hat{g}^{\mathrm{T}}\hat{x} + \hat{c}$

$$\hat{G} = Z^{\mathrm{T}}GZ,\quad \hat{g} = (g + GSb)^{\mathrm{T}} Z,\quad \hat{c} = \frac{1}{2}(2g + GSb)^{\mathrm{T}} Sb.$$

If $\hat{G}$ **PD, the minimizer of QP2:**

$$\hat{x}^* = -\left(Z^{\mathrm{T}}GZ\right)^{-1} Z^{\mathrm{T}}(g + GSb).$$

The minimizer of QPLEC $x^* = Sb - Z\left(Z^{\mathrm{T}}GZ\right)^{-1} Z^{\mathrm{T}}(g + GSb).$

Also find Z and S by QR decomposition of A.

Lagrange Multiplier Method

Let $$L(x,\lambda)=\frac{1}{2}x^{\mathrm{T}}Gx+g^{\mathrm{T}}x-\lambda^{\mathrm{T}}\left(A^{\mathrm{T}}x-b\right).$$

Then Kuhn-Tucker Eq. $$\begin{cases} Gx+g-A\lambda=0 \\ A^{\mathrm{T}}x-b=0 \end{cases}$$

This means that QP can be converted to solution of linear Eq.

Also convert QP to optimization by ASM:

$$\begin{cases} \min\ D(x)=\frac{1}{2}d^{\mathrm{T}}Gd+g^{(k)\mathrm{T}}d, \\ \text{s. t.}\quad a_i^{\mathrm{T}}x=0, i\in I^{(k)}. \end{cases}$$

THANK YOU FOR ATTENDING