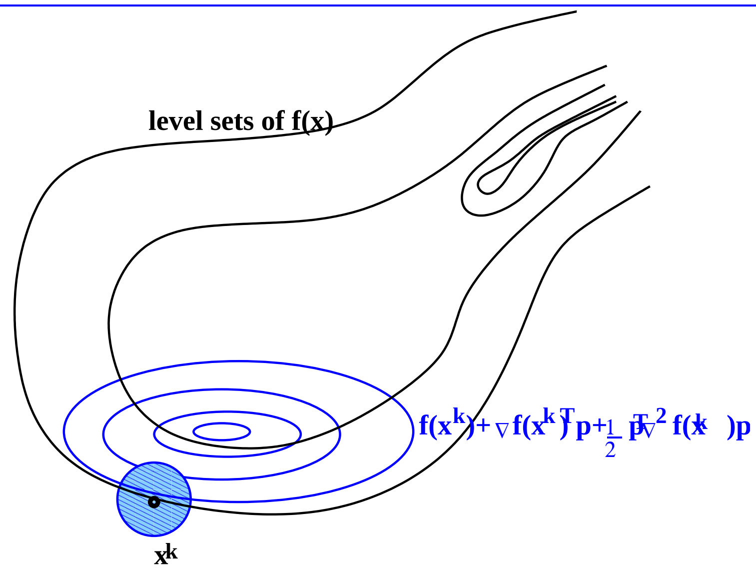
4.4.3　信頼領域法

- 無制約非線形計画問題を考える

$$
\begin{cases}
\text{最小化：} & \boldsymbol{f(x)} \\
\text{条件：} & \boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}
\end{cases}
$$

- $\boldsymbol{k}$ 反復目の点 $\boldsymbol{x^k}$ において信頼半径 $\boldsymbol{\Delta^k}$ として $\boldsymbol{f(x)}$ の $\boldsymbol{2}$ 次近似モデルを考える

$$
\begin{cases}
\text{最小化：} & \boldsymbol{q^k(p) \equiv f(x^k) + \nabla f(x^k)^T p + \frac{1}{2} p^T \nabla^2 f(x^k) p} \\
\text{条件：} & \boldsymbol{\|p\| \le \Delta^k}
\end{cases}
\tag{4.4}
$$



信頼領域法

Step 0: 初期点 $\boldsymbol{x}^0$，初期信頼半径 $\boldsymbol{\Delta}^0$ を与える．パラメータ $0 < \eta_1 \leq \eta_2 < 1,\ 0 < \gamma_1 < 1 < \gamma_2$ を決め，$k = 0$ とする．

終了条件: もし $\boldsymbol{x}^k$ が局所的最適解の近似になっていると判断できたならば終了する．

Step 1: 部分問題 (4.4) を解いて $\boldsymbol{p}^k$ を求める．

Step 2: もし $\frac{f(\boldsymbol{x}^k)-f(\boldsymbol{x}^{k+1})}{q^k(\mathbf{0})-q^k(\boldsymbol{p}^k)=-\nabla f(\boldsymbol{x}^k)^T\boldsymbol{p}^k-\frac{1}{2}(\boldsymbol{p}^k)^T\nabla^2 f(\boldsymbol{x}^k)\boldsymbol{p}^k} \geq \eta_1$ ならば，$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{p}^k$ とする．

- **Step 2.1:** もし $\frac{f(\boldsymbol{x}^k)-f(\boldsymbol{x}^{k+1})}{q^k(\mathbf{0})-q^k(\boldsymbol{p}^k)=-\nabla f(\boldsymbol{x}^k)^T\boldsymbol{p}^k-\frac{1}{2}(\boldsymbol{p}^k)^T\nabla^2 f(\boldsymbol{x}^k)\boldsymbol{p}^k} \geq \eta_2$ ならば，$\boldsymbol{\Delta}^{k+1} \in [\boldsymbol{\Delta}^k, \gamma_2\boldsymbol{\Delta}^k]$ として信頼領域を拡大する．
- **Step 2.2:** そうでなければ $\boldsymbol{\Delta}^{k+1} = \boldsymbol{\Delta}^k$ として現状を維持する．

Step 3: そうでなければ $\boldsymbol{\Delta}^{k+1} \in [\gamma_1\boldsymbol{\Delta}^k, \boldsymbol{\Delta}^k]$ として信頼領域を減少する．

Step 4: $k = k + 1$ とし，終了条件を調べる．

- 部分問題 (4.4) において $\nabla^2 f(\boldsymbol{x}^k)$ を準 Newton 法である BFGS 公式や DFP 公式を用いて更新することが一般的に実装されている

- 部分問題の解法は以下の定理により，固有値問題を解くことによって簡単に計算できる

定理１４

p^kが部分問題(4.4)の最適解である必要十分条件は次の条件を満たしている非負の値$y^k \geq 0$が存在することである．

$$
\begin{aligned}
&(\nabla^2 f(x^k) + y^k I)p^k = -\nabla f(x^k) \\
&y^k(\Delta^k - \|p^k\|) = 0 \\
&(\nabla^2 f(x^k) + y^k I) \text{ は正定値行列}
\end{aligned}
$$

証明

[NOCEDAL2006]などを参照． □

4.4.4 主双対内点法

- 等式制約付き非線形計画問題について考える

$$\begin{cases} \text{最小化：} & \boldsymbol{f(x)} \\ \text{条件：} & \boldsymbol{h_j(x) = 0,\ j = 1, 2, \ldots, m} \\ & \boldsymbol{x \geq 0} \end{cases} \tag{4.5}$$

- 不等式制約 $\boldsymbol{g_i(x) \leq 0,\ i = 1, 2, \ldots, \ell}$ がある場合はスラック変数 $\mathbb{R}^\ell \ni \boldsymbol{z \geq 0}$ を導入して，等式制約 $\boldsymbol{g_i(x) + z_i = 0,\ z_i \geq 0,\ i = 1, 2, \ldots, \ell}$ に変換して上記の問題に帰着することができる

- 問題 (4.5) の KKT 最適性条件は以下のようになる

$$\begin{pmatrix} \boldsymbol{\nabla f(x) + \sum_{j=1}^{m} w_j \nabla h_j(x) - s} \\ \boldsymbol{h(x)} \\ \boldsymbol{XSe} \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{x \geq 0}, \quad \boldsymbol{s \geq 0}$$

ただし，$\boldsymbol{X, S}$ はそれぞれベクトル $\boldsymbol{x, s}$ を対角要素に持つ対角行列であり，$\boldsymbol{e}$ は全ての要素が 1 であるベクトルである．

- 線形計画問題に対する主双対内点法同様にパラメータ $\mu > 0$ を導入して，問題 (4.5) に対する中心パスを以下の方程式の解 $(x(\mu), w(\mu), s(\mu))$ から成る集合だと定義する

$$\boxed{\begin{pmatrix} \nabla f(x) + \sum_{j=1}^{m} w_j \nabla h_j(x) - s \\ h(x) \\ XSe - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad x > 0, \quad s > 0 \quad (4.6)}$$

- 特に $\mu = 0$ の場合，方程式の解は問題 (4.5) の最適解になっている

- ある k 反復目で得られた点 (x^k, w^k, s^k) から次点 $(x^{k+1}, w^{k+1}, s^{k+1})$ を求める際の探索方向 $(\Delta x, \Delta w, \Delta s)$ は上記の方程式の線形化によって求められる．つまり $(x^k + \Delta x, w^k + \Delta w, s^k + \Delta s)$ を線形化した方程式に代入して 2 次項を無視すると以下のようになる

$$\boxed{\begin{array}{l}\begin{pmatrix} \nabla^2 f(x^k) + \sum_{j=1}^{m} w_j^k \nabla^2 h_j(x^k) & \nabla h(x^k) & -I \\ \nabla h^T(x^k) & O & O \\ S^k & O & X^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta w \\ \Delta s \end{pmatrix} \\ = - \begin{pmatrix} \nabla f(x^k) + \sum_{j=1}^{m} w_j^k \nabla h_j(x^k) - s^k \\ h(x^k) \\ X^k S^k e - \mu e \end{pmatrix} \end{array}}$$

- つまり，(4.6) に対する Newton 法の適用になる

- さらに，次点を求めるステップサイズを制御することによって “内点” を保つことができる

$$\begin{aligned} \alpha_p &= \max\{\alpha \in (0,1] : \ x^{k+1} = x^k + \alpha \Delta x \geq 0\} \\ \alpha_d &= \max\{\alpha \in (0,1] : \ s^{k+1} = s^k + \alpha \Delta s \geq 0\} \end{aligned}$$

主双対パス追跡内点法 (実装版)

Step 0: $x^0, s^0 > 0$ となるような初期点 (x^0, w^0, s^0) を決め，$\gamma \in (0,1)$, $k = 0$ とする.

終了条件: もし x^k が近似的に問題 (4.5) の最適解になっているならば終了する

Step 1:

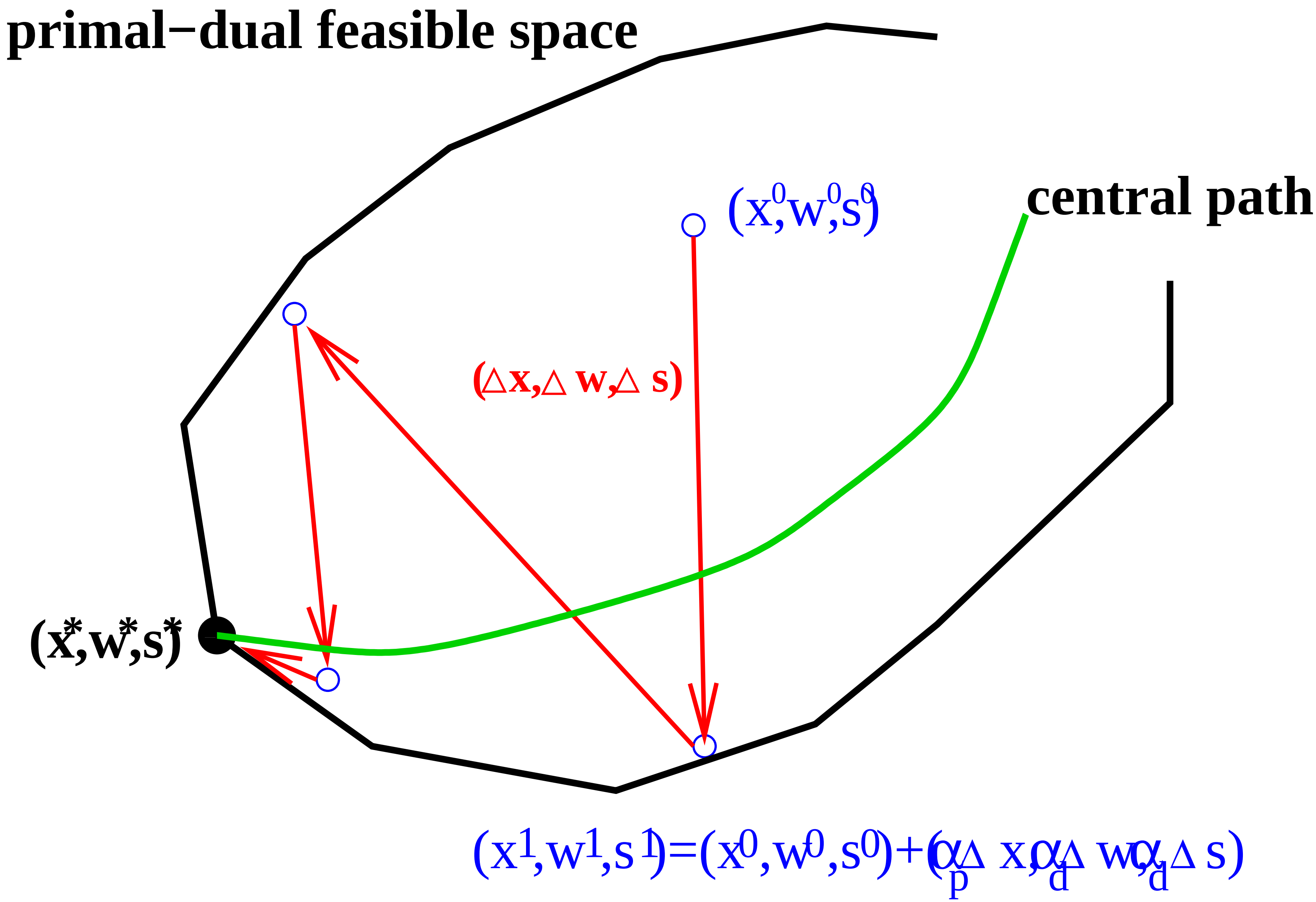
$$\begin{pmatrix} \nabla^2 f(x^k) + \sum_{j=1}^{m} w_j^k \nabla^2 h_j(x^k) & \nabla h(x^k) & -I \\ \nabla h^T(x^k) & O & O \\ S^k & O & X^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta w \\ \Delta s \end{pmatrix} = - \begin{pmatrix} \nabla f(x^k) + \sum_{j=1}^{m} w_j^k \nabla h_j(x^k) - s^k \\ h(x^k) \\ X^k S^k e - \mu e \end{pmatrix}$$

より探索方向 $(\Delta x, \Delta w, \Delta s)$ を求める.

Step 2:

$$\begin{aligned} \alpha_p &= \max\{\alpha \in (0,1] : \ x^{k+1} = x^k + \alpha \Delta x \geq 0\} \\ \alpha_d &= \max\{\alpha \in (0,1] : \ s^{k+1} = s^k + \alpha \Delta s \geq 0\} \end{aligned}$$

Step 3: $(x^{k+1}, w^{k+1}, s^{k+1}) = (x^k + \gamma\alpha_p \Delta x, w^k + \gamma\alpha_d \Delta w, s^k + \gamma\alpha_d \Delta s)$，$k = k+1$ とし，終了条件を調べる.


primal−dual feasible space
central path
(x^0,w^0,s^0)
$(\Delta x, \Delta w, \Delta s)$
(x^*,w^*,s^*)
$(x^1,w^1,s^1)=(x^0,w^0,s^0)+(\alpha_p \Delta x, \alpha_d \Delta w, \alpha_d \Delta s)$

4.4.5 フィルター法

- 一般の非線形計画問題を考える

$$\begin{cases} \text{最小化：} & f(x) \\ \text{条件：} & g_i(x) \leq 0, \ \ i = 1, 2, \ldots, \ell \\ & h_j(x) = 0, \ \ j = 1, 2, \ldots, m \end{cases}$$

- フィルター法では上記の問題にて目的関数 $f(x)$ を減少することと，等式・不等式制約の実行可能性を別々に考える

- 以下の関数によって実行可能性を測る

$$b(x) = \sum_{i=1}^{\ell} |g_i(x)|^+ + \sum_{j=1}^{m} |h_j(x)|$$

ただし，$|g_i(x)|^+ = \max\{0, g_i(x)\}$ である

- フィルター法の目的は以下の 2 つの関数を最小化することにある

最小化 :$f(x)$　　最小化 :$b(x)$

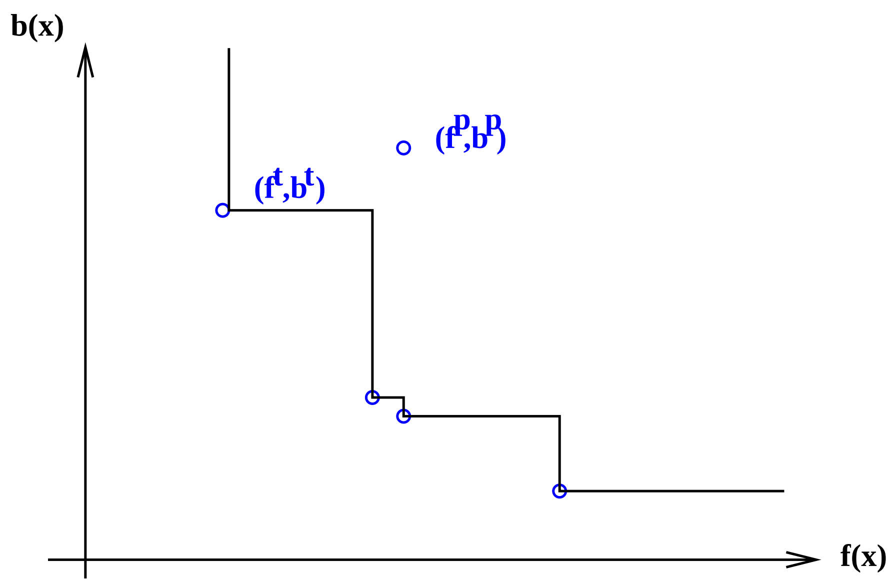
定義15

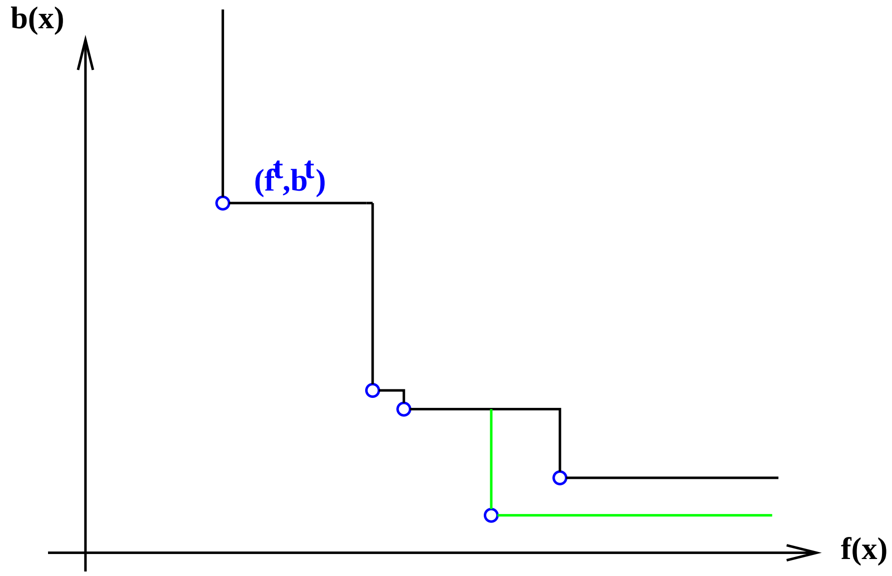
$\boldsymbol{k}$ 反復目の変数 $\boldsymbol{x}^k$ に対して $\boldsymbol{f^k = f(x^k)}$, $\boldsymbol{b^k = b(x^k)}$ と定義する．

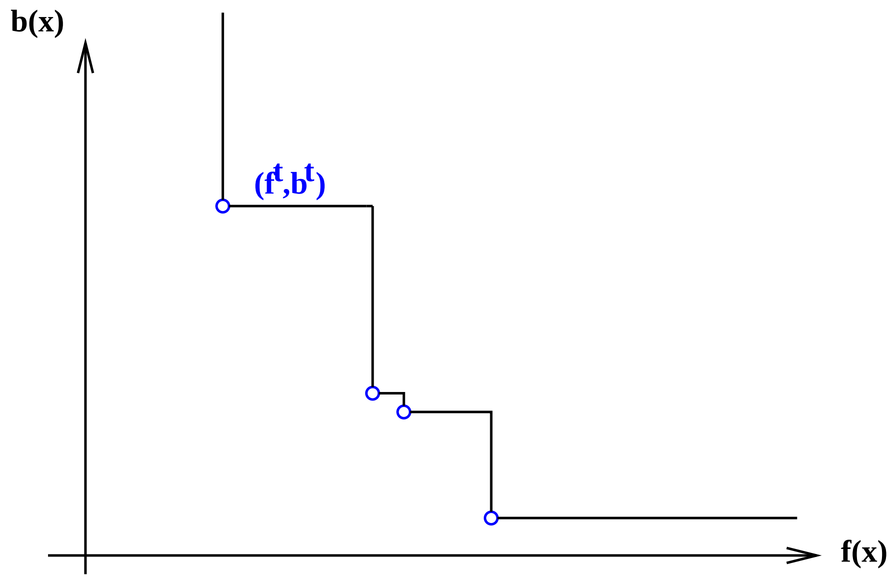
(a) $(\boldsymbol{f}^k, \boldsymbol{b}^k), (\boldsymbol{f}^t, \boldsymbol{b}^t) \in \mathbb{R}^2$ において，$\boldsymbol{f}^k \leq \boldsymbol{f}^t$, $\boldsymbol{b}^k \leq \boldsymbol{b}^t$ が成り立つとき，$(\boldsymbol{f}^k, \boldsymbol{b}^k)$ は $(\boldsymbol{f}^t, \boldsymbol{b}^t)$ より優越しているという．

(b) 自分以外の全てのペアに優越されていないペア $(\boldsymbol{f}^k, \boldsymbol{b}^k)$ から成る集合をフィルターと呼ぶ．

(c) ある点 $\boldsymbol{x}^k$ に対して，$(\boldsymbol{f}^k, \boldsymbol{b}^k)$ がフィルターに入っている全てのペアに優越されていないとき，$\boldsymbol{x}^k$ はフィルターに入るという．



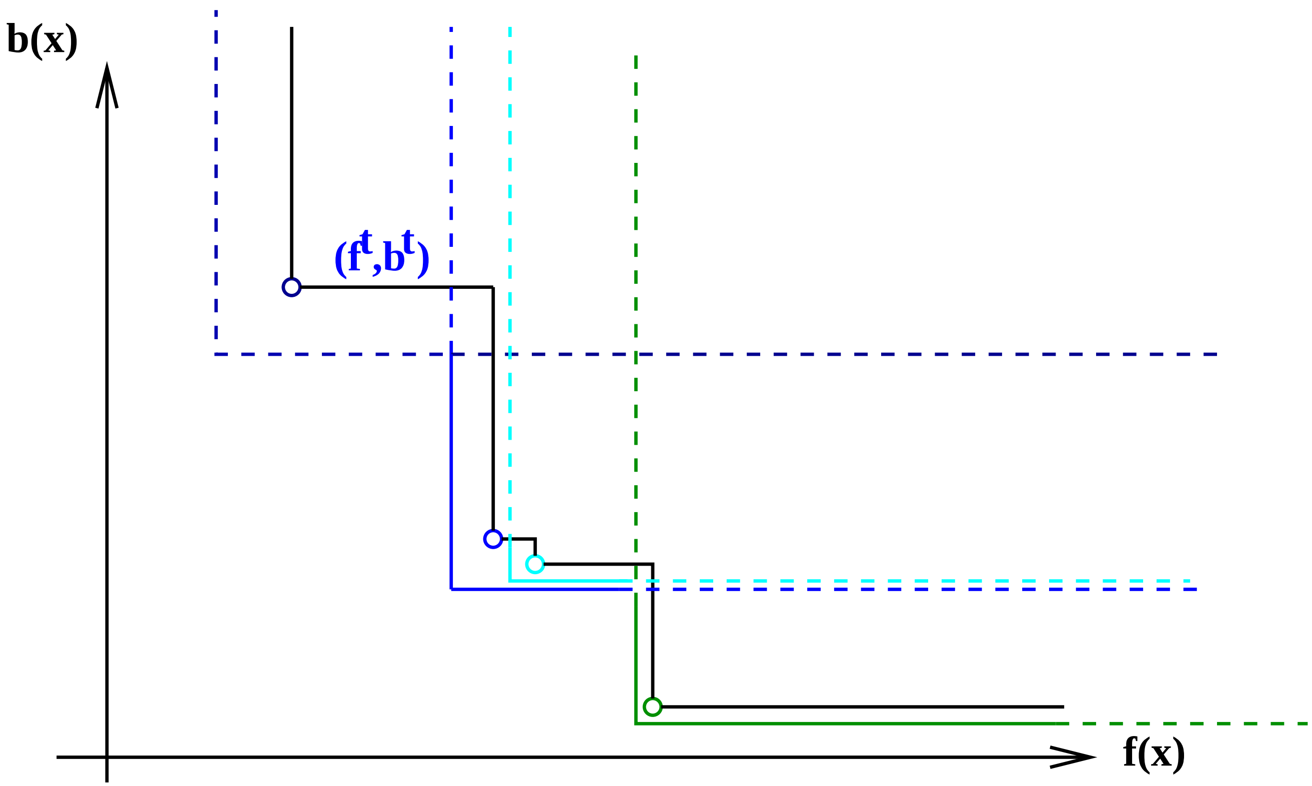

b(x)
(f^t,b^t)
f(x)


b(x)
(f^t,b^t)
f(x)

- 一般の実装において，フィルターと余り近いペアだけが入れられると効率が良くないので，$\beta \in (0,1)$ を用いて，ある禁止区域を導入する．

以下を満たす x^{k+1} に対してのみ (f^{k+1}, b^{k+1}) をフィルターに入れる

$$f(x^{k+1}) \le f^k - \beta b^k \qquad\qquad b(x^{k+1}) \le b^k - \beta b^k$$



フィルター法

Step 0: 初期点 x^0，信頼領域半径 $\Delta^0 > 0$ と $\beta, \alpha_{\min} \in (0,1)$ を決め，$k = 0$ とする．

終了条件: もし x^k が近似的に非線形計画問題の最適解になっているならば終了する．

Step 1: 探索方向 Δx をなにらかの手法で求める．例えば信頼領域もしくは主双対内点法の 1 反復を使って．

Step 2: 直線探索をして (f^{k+1}, b^{k+1}) がフィルターに入るかどうかを調べる．ただし，$(f^{k+1}, b^{k+1}) = (f(x^k + \alpha\Delta x), b(x^k + \alpha\Delta x))$, $\alpha \in (\alpha_{\min}, 1)$ とする．

- **Step 2.1:** もし (f^{k+1}, b^{k+1}) がフィルターに入るならば信頼領域半径を大きくし $\Delta^{k+1} \geq \Delta^k$，$(f^{k+1}, b^{k+1})$ をフィルターに入れて，新しいペアによって優越された全てのペアをフィルターから除く．
- **Step 2.2:** そうでなければ，$x^{k+1} = x^k$ として信頼領域半径を小さくする $\Delta^{k+1} < \Delta^k$.

Step 3: $k = k + 1$ とし，終了条件を調べる．

4.5　演習問題

問題 1

微分可能な関数 $\boldsymbol{f} : \mathbb{R}^n \longrightarrow \mathbb{R}$ が $\mathbb{R}^n$ にて凸関数である必要十分条件は

$$\boldsymbol{f}(\boldsymbol{x}) \geq \boldsymbol{f}(\boldsymbol{y}) + \langle \boldsymbol{\nabla f}(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

であることを示せ．

問題 2

非線形計画問題に対する **Karush-Kuhn-Tucker** の最適性条件とは何か説明せよ．

問題 3

次の線形計画問題

$$\begin{cases} \text{最小化：} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{条件：} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{cases}$$

に対する **Karush-Kuhn-Tucker** 最適性条件を求めよ．

問題 4

次の 2 次計画問題

$$\begin{cases} \text{最小化：} & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} \\ \text{条件：} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{cases}$$

に対するKarush-Kuhn-Tucker最適性条件を求めよ．但し，行列Qは対称行列とみなしてよい．

$\boxed{\text{問題}\,\mathbf{5}}$

勾配ベクトル射影法にて、$\boldsymbol{x}^k = \mathcal{P}(\boldsymbol{x}^k - \nabla\hat{\mathcal{L}}_x(\boldsymbol{x}^k, \boldsymbol{w}^k, \boldsymbol{\mu}^k); \boldsymbol{b}, \boldsymbol{u})$ を満たす$\boldsymbol{x}^k$が該当する最適化問題のKKT条件を満たしていることを示せ．

$\boxed{\text{問題}\,\mathbf{6}}$

線形独立制約想定（定義６）を満たす実行可能解$\boldsymbol{x}^*$はMangasarian-Fromovitzの制約想定（定義１０）を満たすことを示せ．

$\boxed{\text{問題}\,\mathbf{7}}$

非線形計画問題に対する逐次２次近似法について説明せよ．

$\boxed{\text{問題}\,\mathbf{8}}$

信頼領域法について説明せよ．またその実装において，あらかじめどのようなアルゴリズムを用意しておく必要があるか述べよ．

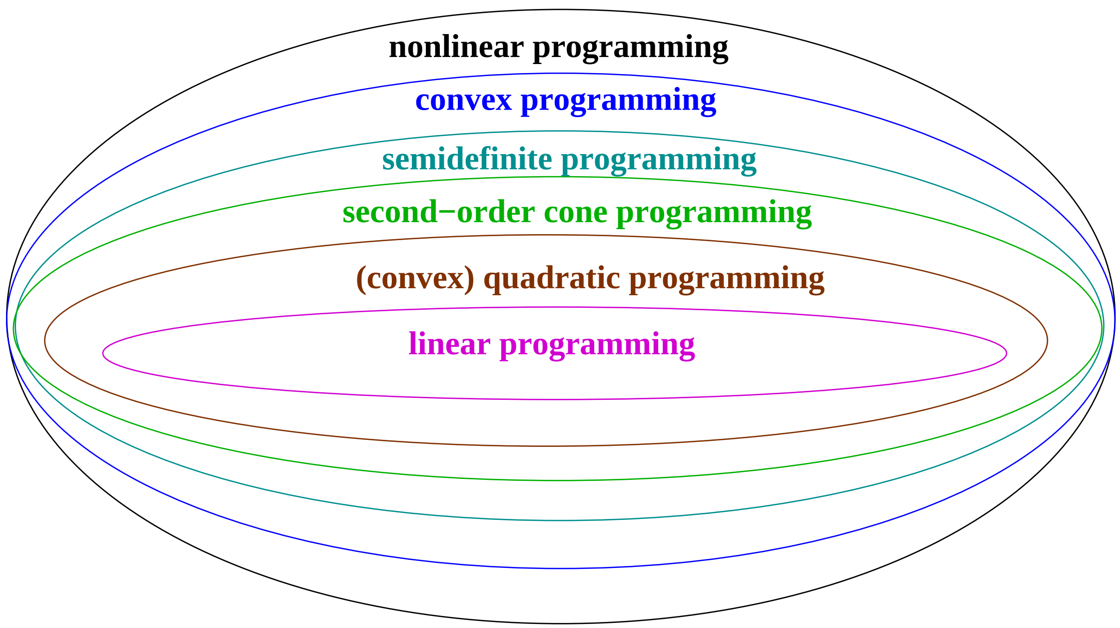
$\boxed{\text{問題}\,\mathbf{9}}$

非線形計画問題に対するフィルター法について説明せよ．

Chapter 5

凸2次計画問題 (Convex Quadratic Programming, CQP)，２次錐計画問題 (Second-Order Cone Programming, SOCP)，半正定値計画問題 (Semidefinite Programming, SDP)

- ここでは非線形計画問題の中でも簡単な問題として分類される2次錐計画問題と半正定値計画問題を扱う
- これらの問題は錐という概念を用いると線形計画問題の拡張としてもみなすことができる
- 多項式時間アルゴリズムである主双対内点法で効率良く解ける
- 線形計画問題より一般的な問題記述能力がある
- 定式化が少し難しいので完全に理解されたり利用されていない



5.1 凸2次計画問題 (Convex Quadratic Programming, CQP)

- まず次の凸2次計画問題について考える

$$\begin{cases} \text{最小化：} & \frac{1}{2}x^TQx + q^Tx \\ \text{条件：} & Ax = b \\ & x \geq 0 \end{cases} \tag{5.1}$$

ただし，$Q \succeq O$

- ここで $f(x) = \frac{1}{2}x^TQx + q^Tx$ は凸関数である
- これは $Q \succeq O$，つまり対称行列 Q が半正定値行列であることと同値である
- $Q \succeq O \Leftrightarrow Q$ の全ての (実数) 固有値は非負である

5.1.1 応用例：Markovitzのポートフォリオモデル

- n株の投資を考え，各銘柄iに投資する比率x_iを求めたい

株iの収益率	r_i	$i = 1, 2, \ldots, n$ (正規分布に従う確率変数)
株iの収益率の期待値	$\mu_i = E[r_i]$	$i = 1, 2, \ldots, n$
株iの収益率の分散	$\sigma_i^2 = E[(r_i - \mu_i)^2]$	$i = 1, 2, \ldots, n$
株iとjの共分散	$E[(r_i - \mu_i)(r_j - \mu_j)]$	$i, j = 1, 2, \ldots, n$
株iとjの相関係数	$\rho_{i,j} = \frac{E[(r_i - \mu_i)(r_j - \mu_j)]}{\sigma_i \sigma_j}$	$i, j = 1, 2, \ldots, n$
各株に投資する比率	$\mathbb{R}^n \ni x \geq 0$	$\sum_{i=1}^{n} x_i = 1$
全体の収益率	$\sum_{i=1}^{n} x_i \mu_i = \mu^T x$	
全体の分散	$\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_i \sigma_j \rho_{i,j} x_i x_j = x^T G x$	$G \succeq O$は分散共分散行列

- 全体の収益率と分散リスクのバランスを考慮するパラメータ$\kappa \geq 0$を導入する

$$
\left\{
\begin{array}{ll}
\text{最大化：} & \mu^T x - \kappa x^T G x \\
\text{条件：} & e^T x = 1 \\
 & x \geq 0
\end{array}
\right.
$$

$$
\begin{cases}
\text{最大化：} & \boldsymbol{\mu}^T \boldsymbol{x} - \kappa \boldsymbol{x}^T G \boldsymbol{x} \\
\text{条件：} & \boldsymbol{e}^T \boldsymbol{x} = 1 \\
& \boldsymbol{x} \geq 0
\end{cases}
$$

- 保守的な選択でリスクを重視する場合は κ を大きくあらかじめ決める
- 利益を優先する場合は κ を小さく決める
- 分散共分散行列 G は半正定値行列である
- 目的関数の符号を反転することにより (5.1) に帰着し、凸 2 次計画問題になる

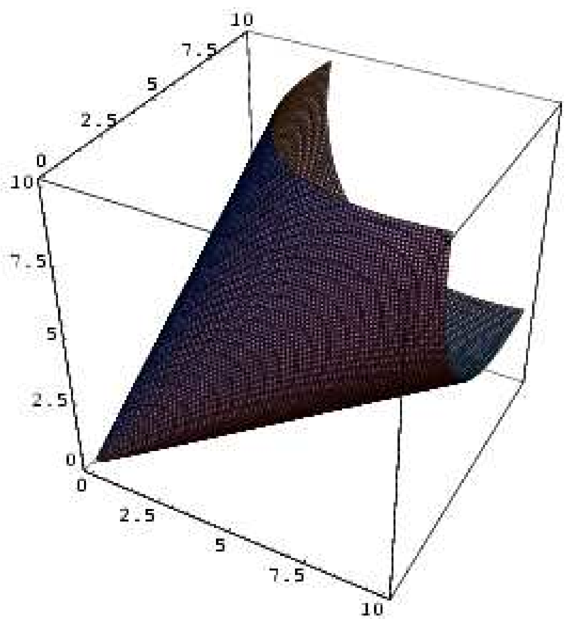
$$
\begin{cases}
\text{最小化：} & \kappa \boldsymbol{x}^T G \boldsymbol{x} - \boldsymbol{\mu}^T \boldsymbol{x} \\
\text{条件：} & \boldsymbol{e}^T \boldsymbol{x} = 1 \\
& \boldsymbol{x} \geq 0
\end{cases}
$$

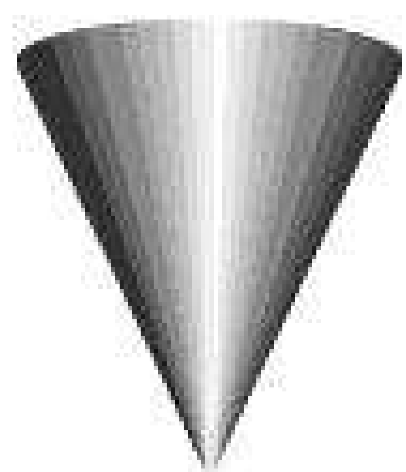
- (5.1) は単体法，内点法や特化されたアルゴリズムによって解ける．

5.2 2次錐計画問題 (Second-Order Cone Programming, SOCP)

2次錐

$$\mathcal{K}^n = \left\{ (x_0; z) \in \mathbb{R} \times \mathbb{R}^{n-1} : \sqrt{\sum_{i=1}^{n-1} z_i^2} \le x_0 \right\}$$
$$= \left\{ (x_0; z) \in \mathbb{R} \times \mathbb{R}^{n-1} : \|z\|_2 \le x_0 \right\}$$





定義 1

ある集合$\boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^\ell$が2次錐制約で書き表わせるとは$\mathbb{R}^\ell$から$\mathbb{R}^n$へのアフィン写像が存在し，その値域が2次錐に入っている時をいう．

つまり，行列$\boldsymbol{A} \in \mathbb{R}^{n\times\ell}$とベクトル$\boldsymbol{b} \in \mathbb{R}^n$が存在し，

$$\boldsymbol{y} \in \boldsymbol{\mathcal{C}} \subseteq \mathbb{R}^\ell \;\Leftrightarrow\; (\boldsymbol{x}_0, \boldsymbol{z}) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}, \quad (\boldsymbol{x}_0, \boldsymbol{z}) \in \boldsymbol{\mathcal{K}}^n \subseteq \mathbb{R}^n$$

となる時である．

この場合，集合$\boldsymbol{\mathcal{C}}$は凸集合である．

命題２

ベクトル $\boldsymbol{w} \in \mathbb{R}^n$ とスカラー $\boldsymbol{t}, \boldsymbol{s} \in \mathbb{R}$ についての制約

$$\boldsymbol{w}^T\boldsymbol{w} \le \boldsymbol{ts}, \qquad \boldsymbol{t} \ge \boldsymbol{0}, \quad \boldsymbol{s} \ge \boldsymbol{0}$$

(双曲型制約) は２次錐制約を用いて

$$\left\| \begin{pmatrix} \boldsymbol{t} - \boldsymbol{s} \\ \boldsymbol{2w} \end{pmatrix} \right\|_2 \le \boldsymbol{t} + \boldsymbol{s}$$

として書ける．つまり，$(\boldsymbol{t} + \boldsymbol{s}; (\boldsymbol{t} - \boldsymbol{s}, \boldsymbol{2w})) \in \mathcal{K}^{n+2}$.

証明

$$4\boldsymbol{w}^T\boldsymbol{w} + \boldsymbol{t}^2 + \boldsymbol{s}^2 \le 4\boldsymbol{ts} + \boldsymbol{t}^2 + \boldsymbol{s}^2 \Leftrightarrow \sum_{i=1}^n (2\boldsymbol{w}_i)^2 + (\boldsymbol{t} - \boldsymbol{s})^2 \le (\boldsymbol{t} + \boldsymbol{s})^2$$

$$\Leftrightarrow \left\| \begin{pmatrix} \boldsymbol{t} - \boldsymbol{s} \\ \boldsymbol{2w} \end{pmatrix} \right\|_2 \le \boldsymbol{t} + \boldsymbol{s}$$

□

以下，２次錐制約で書ける例題を幾つか示す

5.2.1 凸2次制約

- 次の凸2次制約を考える：$\frac{1}{2}x^TQx + q^Tx + p \le 0$，ただし，$Q \succeq O$は対称半正定値行列, $q \in \mathbb{R}^n$, $p \in \mathbb{R}$はスカラーである．
- Qは半正定値行列なので $Q = R^TR$となるような行列$R \in \mathbb{R}^{n\times n}$ が存在する

$$\begin{aligned}
\frac{1}{2}x^TQx + q^Tx + p &\le 0 \\
\frac{1}{4}x^TR^TRx &\le -\frac{1}{2}(q^Tx+p) \\
\left(\frac{1}{2}Rx\right)^T\left(\frac{1}{2}Rx\right) &\le -\frac{1}{2}(q^Tx+p)1
\end{aligned}$$

命題2において$w = \frac{1}{2}Rx$, $t = -\frac{1}{2}(q^Tx+p)$, $s = 1$と置き，上記の左辺は常に非負であることから$t, s \ge 0$になるので

$$\left\| \begin{array}{c} -\frac{1}{2}(q^Tx+p) - 1 \\ Rx \end{array} \right\| \le -\frac{1}{2}(q^Tx+p) + 1$$

$$\left(-\frac{1}{2}(q^Tx+p)+1; \left(\begin{array}{c} -\frac{1}{2}(q^Tx+p) - 1 \\ Rx \end{array}\right)\right) \in \mathcal{K}^{n+2}$$