

$$V_k = I - \frac{s^k (y^k)^T}{(y^k)^T s^k}$$

とおくと BFGS 公式 (3.3) は

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

と書き換えることができる．さらにこれを逐次的に代入していくと次の式が得られる．

$$\begin{aligned} H_k = & (V_0 \cdots V_{k-1})^T H_0 (V_0 \cdots V_{k-1}) + (V_1 \cdots V_{k-1})^T \frac{y^0 (y^0)^T}{(y^0)^T s^0} (V_1 \cdots V_{k-1}) \\ & + \cdots + V_{k-1}^T \frac{y^{k-2} (y^{k-2})^T}{(y^{k-2})^T s^{k-2}} V_{k-1} + \frac{y^{k-1} (y^{k-1})^T}{(y^{k-1})^T s^{k-1}} \end{aligned}$$

反復が進むに連れ、最適解に近づくことを考えると、初期の反復から得た情報をある程度無視してもよい気がする．それで、最新の q 反復 の情報 $V_{k-1}, \dots, V_{k-q}$ だけを考えて、上記の式は次のようになる．

$$\begin{aligned}\widetilde{H}_k &= (V_{k-q} \cdots V_{k-1})^T H_k^0 (V_{k-q} \cdots V_{k-1}) \\ &\quad + (V_{k-q+1} \cdots V_{k-1})^T \frac{y^{k-q} (y^{k-q})^T}{(s^{k-q})^T y^{k-q}} (V_{k-q+1} \cdots V_{k-1}) \\ &\quad + \cdots + V_{k-1}^T \frac{y^{k-2} (y^{k-2})^T}{(y^{k-2})^T s^{k-2}} V_{k-1} + \frac{y^{k-1} (y^{k-1})^T}{(y^{k-1})^T s^{k-1}}\end{aligned}$$

ただし， H_k^0 は疎性を持った行列とする．

さらに，実質的には

$$\boxed{H_k \nabla f(x^k)} \quad \text{もしくは} \quad \boxed{\widetilde{H}_k \nabla f(x^k)}$$

が必要なので H_k もしくは $\widetilde{H}_k$ において右からベクトル $\nabla f(x^k)$ をかけるとほとんど全ての演算がベクトルの内積で出来ることがわかる．

＜記憶制限準ニュートン法 (L-BFGS) : アルゴリズム＞

Step 0: x^0 を初期点, $q > 0$ を遡る反復回数, $\epsilon > 0$ を停止基準, $H_0 := I$, $k := 0$ を反復回数とする.

Step 0': H_k^0 を与える (例えば疎行列 $((y^{k-1})^T s^{k-1}) / ((s^{k-1})^T s^{k-1}) I$ とする).

Step 1: $k \leq q$ の場合は $d^k := -H_k \nabla f(x^k)$, そうでなければ $d^k := -\widetilde{H}_k \nabla f(x^k)$ を計算する.

Step 2: $\alpha \geq 0$ に対する以下の 直線探索 を近似的に解く.

$$\{ \text{最小化: } f(x^k + \alpha d^k); \text{ 条件: } \alpha \geq 0$$

この解を α^k として, $x^{k+1} := x^k + \alpha^k d^k$ とおく.

Step 3: $\|x^{k+1} - x^k\| < \epsilon$ ならば終了. y^k, s^k を求める.

Step 4: $k := k + 1$ として, Step 0' へ.

- 特に, $q = k$ の場合, 記憶制限準ニュートン法は準ニュートン法になる.
- $q = 1$ の場合, 次に紹介する共役勾配法になる.

<利点>

- ヘッセ行列やその逆行列を必要としないし，保持する必要もない．
- 大規模な問題にも適用できる．

<欠点>

- ニュートン法よりも数値的な安定性に欠ける．特にヘッセ行列の固有値が幅広く分布されている場合など．

3.11 共役勾配法

<基本的な考え方>

- ニュートン法や準ニュートン法では陽にヘッセ行列を保持する必要があるため，大規模な問題が解けない．
- ヘッセ行列やその逆行列を保持しないで行列ベクトル乗算のみで済ませる．

＜共役勾配法：アルゴリズム＞

Step 0: x^0 を初期点, $f(x^0)$, $\nabla f(x^0)$ を計算し, $p^0 := -\nabla f(x^0)$, $\epsilon > 0$ を停止基準, $k := 0$ を反復回数とする.

Step 1: $\alpha \geq 0$ に対する以下の直線探索を近似的に解く.

$$\{ \text{最小化: } f(x^k + \alpha p^k); \text{ 条件: } \alpha \geq 0$$

この解を α^k として, $x^{k+1} := x^k + \alpha^k p^k$ とおく.

Step 2: $\|x^{k+1} - x^k\| < \epsilon$ ならば終了. そうでなければ $k := k + 1$ とする.

Step 3: $f(x^k)$, $\nabla f(x^k)$ を計算する.

Step 4: β_k を次の公式から求める.

Step 4: $p^k := -\nabla f(x^k) + \beta_k p^{k-1}$ として Step 1 へ.

Hestenes-Stiefel 公式 (1952)

$$\beta_k := \frac{\nabla f(x^k)^T (\nabla f(x^k) - \nabla f(x^{k-1}))}{(p^{k-1})^T (\nabla f(x^k) - \nabla f(x^{k-1}))}$$

Fletcher-Reeves 公式 (1964)

$$\beta_k := \frac{\|\nabla f(x^k)\|^2}{\|\nabla f(x^{k-1})\|^2}$$

Polak-Ribière 公式

$$\beta_k := \frac{\nabla f(x^k)^T (\nabla f(x^k) - \nabla f(x^{k-1}))}{\|\nabla f(x^{k-1})\|^2}$$

Polak-Ribière plus 公式

$$\beta_k := \max \left\{ 0, \frac{\nabla f(x^k)^T (\nabla f(x^k) - \nabla f(x^{k-1}))}{\|\nabla f(x^{k-1})\|^2} \right\}$$

Dai-Yuan 公式 (1999)

$$\beta_k := \frac{\|\nabla f(x^k)^T\|^2}{(p^{k-1})^T (\nabla f(x^k) - \nabla f(x^{k-1}))}$$

<利点>

- ヘッセ行列を必要としないし，その保持の必要もない．

- 狭義凸 2 次関数に対しては有限回の反復で収束する.

<欠点>

- ニュートン法よりも数値的な安定性に欠ける.

3.12 演習問題

問題 1

黄金分割法による直線探索の利点を述べよ.

問題 2

2 次関数補間法にて, 2 点 $\alpha_1, \alpha_2 \in \mathbb{R}$ が与えられていて, 関数 $\phi(\alpha)$ を 2 次関数 $q(\alpha) = a\alpha^2 + b\alpha + c$ で補間するとする. この時, $q(\alpha_1) = \phi(\alpha_1)$, $q(\alpha_2) = \phi(\alpha_2)$, $q'(\alpha_1) = \phi'(\alpha_1)$ となるように, 関数 $q(\alpha)$ の最小解 $\bar{\alpha}$ を求めよ.

問題 3

2 次関数補間法にて, 2 点 $\alpha_1, \alpha_2 \in \mathbb{R}$ が与えられていて, 関数 $\phi(\alpha)$ を 2 次関数 $q(\alpha) = a\alpha^2 + b\alpha + c$ で補間するとする. この時, $q(\alpha_1) = \phi(\alpha_1)$, $q'(\alpha_1) = \phi'(\alpha_1)$, $q'(\alpha_2) = \phi'(\alpha_2)$ となるように, 関数 $q(\alpha)$ の最小解 $\bar{\alpha}$ を求めよ.

問題 4

非線形実数値関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ が与えられたとき, $x^0 = 0$ と f の極小点を含む区間を求めたい (いわゆる “囲い込み”). どのようにすればよいかを説明せよ.

問題 5

制約条件の付かない非線形連続最適化問題に対するシンプレックス法とパターン探索法を説明よせ.

問題 6

定理 4 の結果から $x^0 \in \mathbb{R}^n$ を初期点とした厳密直線探索を用いた最急降下法が $f(x)$ に適用された時, $\|x^k - x^*\| < \epsilon$ を得るには反復回数 k が $\mathcal{O}(\ln(1/\epsilon))$ 必要であることを示せ.

問題 7

定理 7 の仮定のもと, $x^0 \in \mathbb{R}^n$ を初期点としてニュートン法を適用して求まる点列にて, $\|x^k - x^*\| < \epsilon$ を得るには反復回数 k が $\mathcal{O}(\ln \ln(1/\epsilon))$ 必要であることを示せ.

問題 8

制約条件の付かない非線形連続最適化問題に対する最急降下法と Newton 法について説明し, それぞれの利点および欠点をあげよ.

問題 9

凸 2 次計画問題を最小化する際, 加速勾配法が最急下法に勝つと思われる理由を記せ.

問題 10

準ニュートン法において, 対称ランク 1 公式, DFP 公式, BFGS 公式がそれぞれ $H_{k+1}s^k =$

y^k を満たすことを示せ.

問題 11

$u, v \in \mathbb{R}^n$, 正則行列 $M \in \mathbb{R}^{n \times n}$ に対して, $1 + v^T M^{-1} u \neq 0$ ならば,

$$(M + uv^T)^{-1} = M^{-1} - \frac{M^{-1}uv^T M^{-1}}{1 + v^T M^{-1}u}$$

(Sherman-Morrison 公式) が成り立つ. この公式を使い, 対称ランク 1 公式, DFP 公式, BFGS 公式から H_{k+1} の逆行列をそれぞれ求めよ.

問題 12

$f(x) = x^T Q x / 2 + q^T x$ として定義された狭義凸 2 次関数を考える. ただし, Q は正定値対称行列とする. $f(x)$ に共役勾配法を適用した時, Fletcher-Reeves 公式, Polak-Ribière 公式, Hestenes-Stiefel 公式, Dai-Yuan 公式から求まる係数 β_k が一致することを示せ.

問題 13

再度, 共役勾配法を狭義凸 2 次関数 $f(x) = x^T Q x / 2 + q^T x$ に対して適用することを考える. ただし, Q は正定値対称行列とする. この時, 共役勾配法で生成される p^k ($k = 0, 1, \dots$) が $(p^i)^T Q p^j = 0$ ($i, j = 0, 1, \dots$) を満たすことを示せ.

Chapter 4

非線形計画問題, Nonlinear Programming, NLP

4.1 定式化, 用語

(制約付き) 非線形計画問題

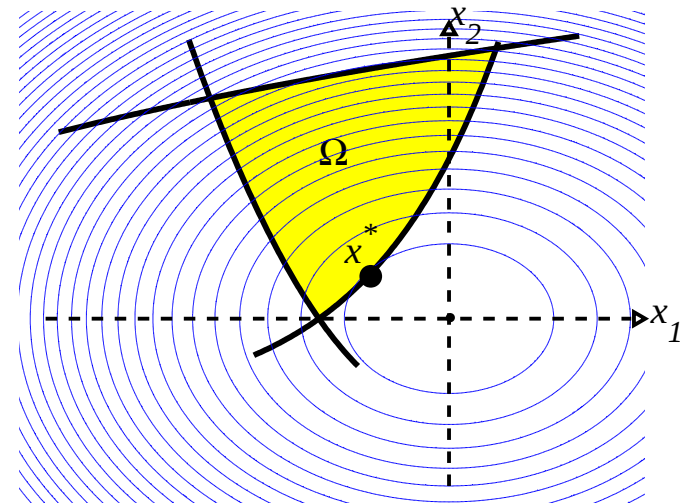
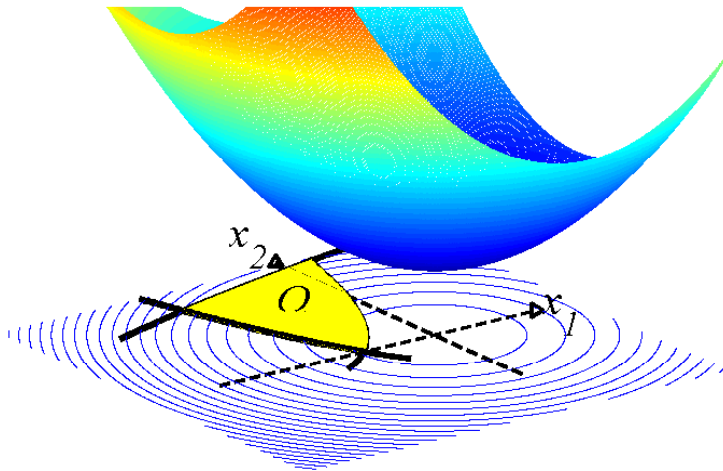
$$\left\{ \begin{array}{l} \text{最小化: } f(x) \\ \text{条件: } \quad g_i(x) \leq 0, \quad i = 1, 2, \dots, \ell \\ \quad \quad h_j(x) = 0, \quad j = 1, 2, \dots, m \end{array} \right.$$

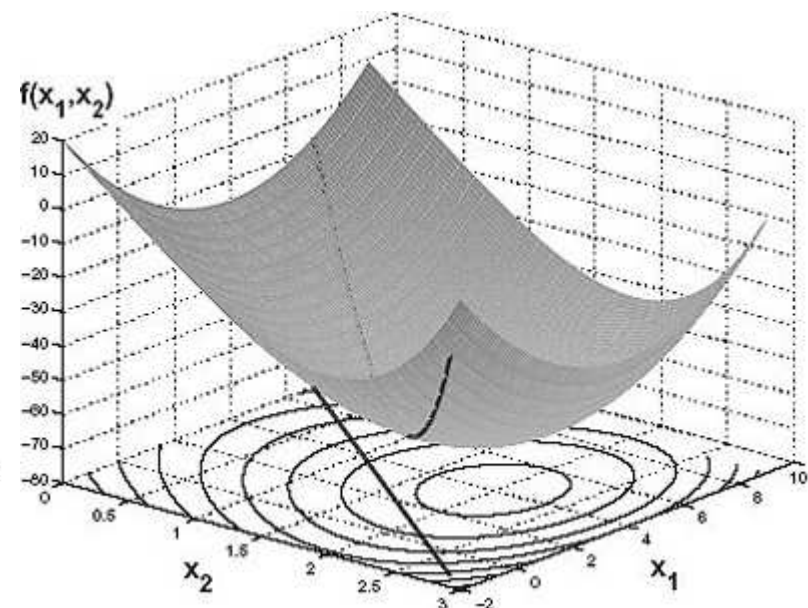
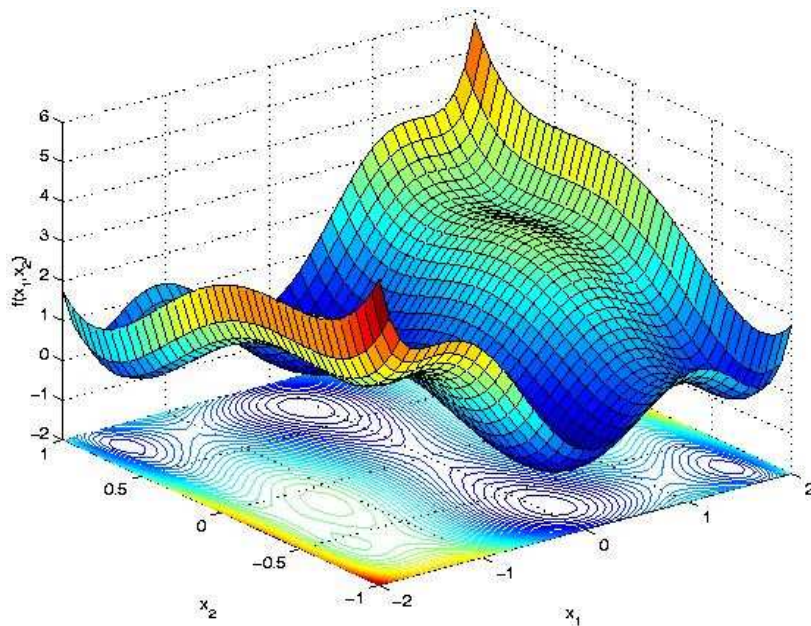
ただし, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h_j : \mathbb{R}^n \longrightarrow \mathbb{R}$

(制約付き) 非線形計画問題

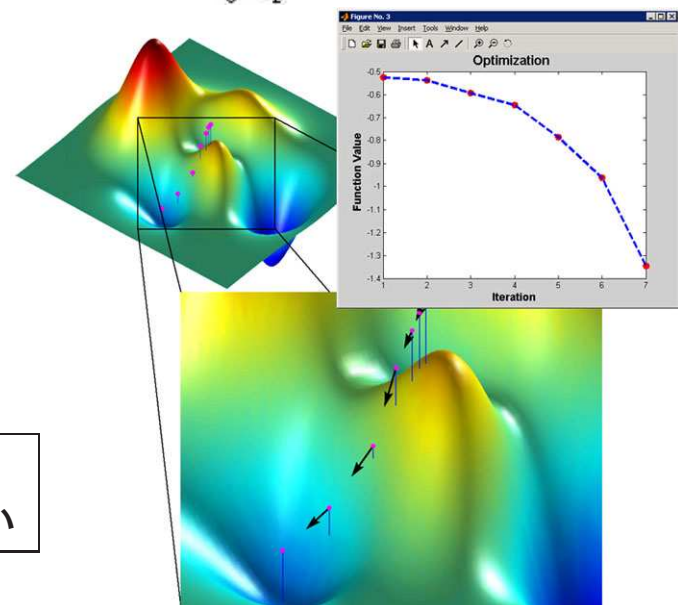
$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } g_i(x) \leq 0, \quad i = 1, 2, \dots, \ell \\ \quad \quad h_j(x) = 0, \quad j = 1, 2, \dots, m \end{cases}$$

ただし, $f: \mathbb{R}^n \longrightarrow \mathbb{R}$, $g_i: \mathbb{R}^n \longrightarrow \mathbb{R}$, $h_j: \mathbb{R}^n \longrightarrow \mathbb{R}$





- 大域的最適解と大域的最適値 $\Rightarrow$ 比較的難しい
- 局所的最適解と局所的最適値 $\Rightarrow$ 変数が多いと
局所的な情報しかない



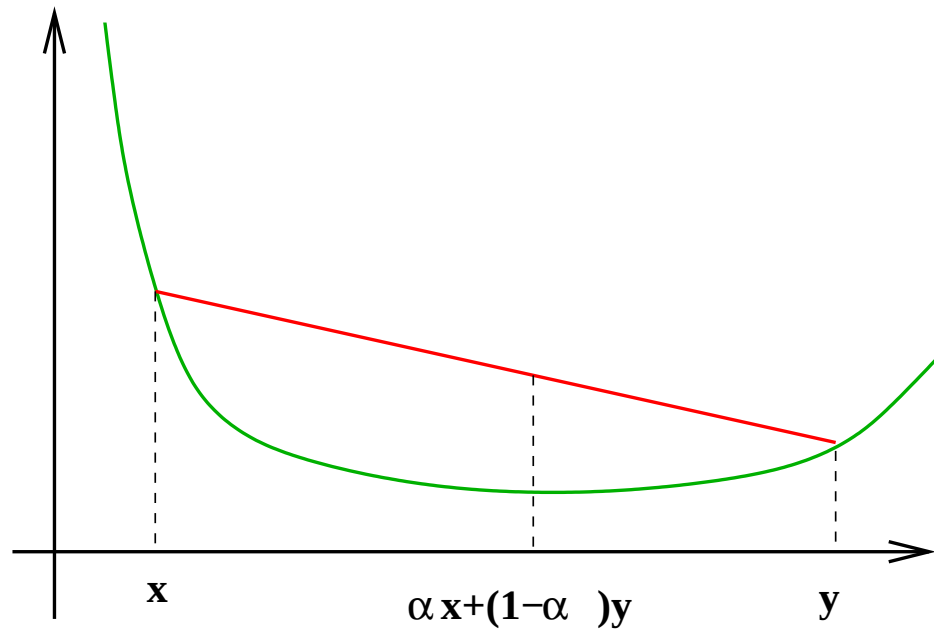
4.2 凸関数

定義 1

関数 $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ は以下の条件を満たしていれば **凸関数** と呼ばれる.

$$\forall x, y \in \mathbb{R}^n, \quad \forall \alpha \in [0, 1]$$

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

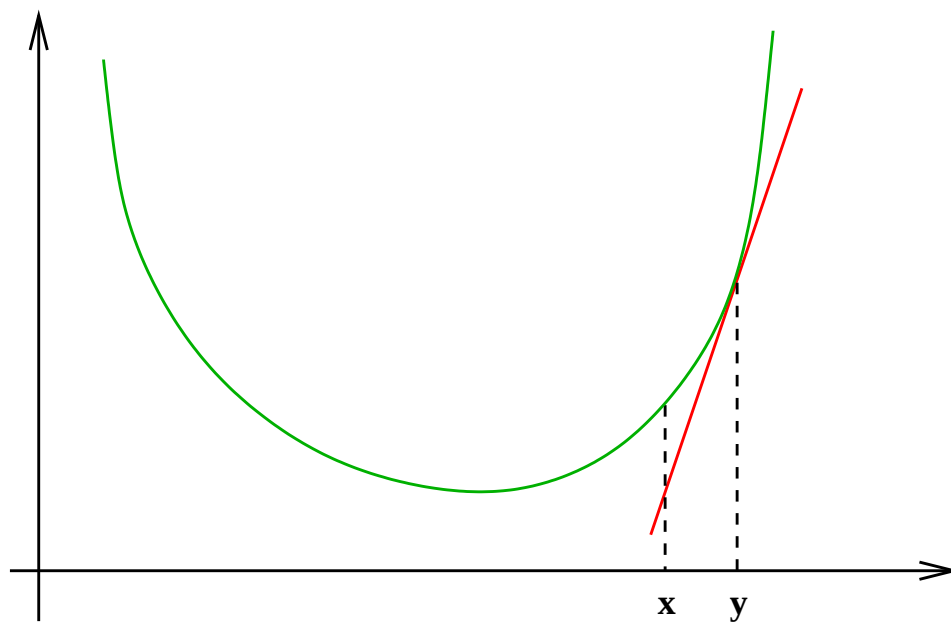


命題 2

定義域 $\mathbb{R}^n$ で定義された 微分可能 な関数 f が凸関数である必要十分条件は

$$\forall x, y \in \mathbb{R}^n$$

$$f(x) \geq f(y) + \langle \nabla f(y), (x - y) \rangle$$



4.3 最適性条件

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

仮定

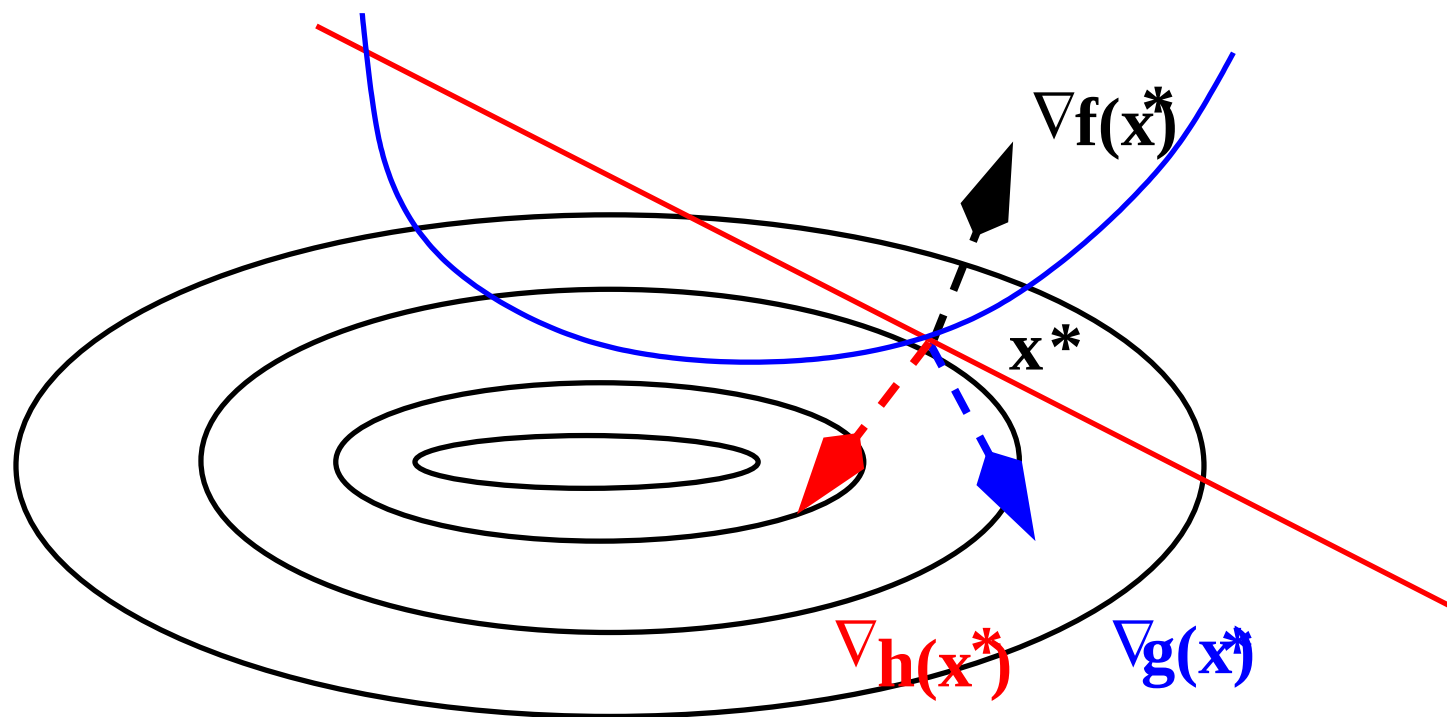
f, g_i, h_j は定義域 $\mathbb{R}^n$ にて
微分可能な関数である

Karush-Kuhn-Tucker の最適性条件

次の条件を満たす $x^* \in \mathbb{R}^n$ と $(y^*, w^*) \in \mathbb{R}^\ell \times \mathbb{R}^m$ が存在する

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{\ell} y_i^* \nabla g_i(x^*) + \sum_{j=1}^m w_j^* \nabla h_j(x^*) &= 0 \\ g_i(x^*) &\leq 0, \quad i = 1, 2, \dots, \ell \\ h_j(x^*) &= 0, \quad j = 1, 2, \dots, m \\ y_i^* g_i(x^*) &= 0, \quad i = 1, 2, \dots, \ell \\ y^* &\geq 0 \end{aligned} \tag{4.1}$$

例 1



例 2

$$\left\{ \begin{array}{l} \text{最小化: } x_1^2 + x_2^2 \\ \text{条件: } \quad x_1^2 + x_2^2 - 5 \leq 0 \\ \quad \quad -x_1 \leq 0 \\ \quad \quad -x_2 \leq 0 \\ \quad \quad x_1 + 2x_2 - 4 = 0 \end{array} \right.$$

$x^* = \left(\frac{4}{5}, \frac{8}{5} \right)$ は最適解であり, $y^* = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $w^* = -\frac{8}{5}$ ととると, (x^*, y^*, w^*) は KKT

の最適性条件を満たす.

(制約付き) 非線形計画問題

$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } g_i(x) \leq 0, \quad i = 1, 2, \dots, \ell \\ \quad \quad h_j(x) = 0, \quad j = 1, 2, \dots, m \end{cases} \quad (4.2)$$

x^* が(4.2)の局所的最適解である

必要条件 $\Downarrow$ $\Uparrow$ 十分条件

次の条件を満たす (y^*, w^*) が存在する

Karush-Kuhn-Tucker の最適性条件

$$\nabla f(x^*) + \sum_{i=1}^{\ell} y_i^* \nabla g_i(x^*) + \sum_{j=1}^m w_j^* \nabla h_j(x^*) = 0$$

$$\begin{aligned} g_i(x^*) &\leq 0, \quad i = 1, 2, \dots, \ell \\ h_j(x^*) &= 0, \quad j = 1, 2, \dots, m \\ y_i^* g_i(x^*) &= 0, \quad i = 1, 2, \dots, \ell \\ y^* &\geq 0 \end{aligned}$$

4.3.1 最適性十分条件

定理 3 (Karush-Kuhn-Tucker の最適性十分条件)

$x^*, (y^*, w^*)$ が Karush-Kuhn-Tucker の最適性条件を満たしているとする. $f, g_i, i = 1, 2, \dots, \ell$ が凸関数であり, $h_j, j = 1, 2, \dots, m$ が線形関数だとすると, x^* は非線形計画問題 (4.2) の局所的最適解になる.

証明

[矢部 2006] などを参照.

□

定義 4

非線形計画問題 (4.2) の不等式制約 $g_i, i = 1, 2, \dots, \ell$ とある点 x^* に対して, $g_k(x^*) = 0$ が成り立つとき, この制約式は点 x^* で効いているといい, この制約式は x^* で有効制約式であるという. また $g_k(x^*) < 0$ のとき, この制約式は点 x^* で効いていないという.

x^* に対して有効な不等式制約の添字集合を $I(x^*)$ で表わす.

例 2 において $x^* = (\frac{4}{5}, \frac{8}{5})^T$ は $x_1^2 + x_2^2 - 5 \leq 0, -x_1 \leq 0, -x_2 \leq 0$ に対して有効制約式でない.

定理 5 (2次の最適性十分条件)

x^* , (y^*, w^*) が Karush-Kuhn-Tucker の最適性条件を満たしているとする．さらに不等式制約の中で有効制約式 $g_i(x^*) = 0$, $i \in I(x^*)$ に対して $y_i^* > 0$ であり, かつ

$$u^T \left(\nabla^2 f(x^*) + \sum_{i \in I(x^*)} y_i^* \nabla^2 g_i(x^*) + \sum_{j=1}^m w_j^* \nabla^2 h_j(x^*) \right) u > 0,$$

が

$$\nabla g_i(x^*)^T u = 0, \quad i \in I(x^*), \quad \nabla h_j(x^*)^T u = 0, \quad j = 1, 2, \dots, m$$

を満たす全ての $u \neq 0$ に対していえるのならば, x^* は非線形計画問題 (4.2) の狭義局所的最適解である．

証明

[BERTSEKAS2003] など参照.

□

4.3.2 最適性必要条件

定義 6

非線形計画問題 (4.2) の実行可能解 x^* , つまり不等式制約にて $g_i(x^*) \leq 0$, $i = 1, 2, \dots, \ell$, 等式制約にて $h_j(x^*) = 0$, $j = 1, 2, \dots, m$ を満たす点が線形独立制約想定を満たしているとは不等式制約の中で有効な制約式の勾配ベクトル $\nabla g_i(x^*)$, $i \in I(x^*)$ と等式制約の勾配ベクトル $\nabla h_j(x^*)$, $j = 1, 2, \dots, m$ が線形独立であるときをいう.

例 3

$$\left\{ \begin{array}{ll} \text{最小化:} & f(x_1, x_2, x_3) \\ \text{条件:} & \begin{array}{ll} 3x_1^2 - x_2 + x_3^3 - 1 & \leq 0 \\ -x_1 - \frac{3}{8} & \leq 0 \\ x_2 - \frac{27}{64} & \leq 0 \\ 3x_1 + 4x_2 - \frac{9}{16} & = 0 \end{array} \end{array} \right.$$

$\begin{pmatrix} \frac{1}{8} \\ \frac{3}{64} \\ 1 \end{pmatrix}$ は線形独立制約想定を満たしているが, $\begin{pmatrix} -\frac{3}{8} \\ \frac{27}{64} \\ 0 \end{pmatrix}$ は線形独立制約想定を満たしていない.

定義 6 (再掲)

非線形計画問題 (4.2) の実行可能解 x^* , つまり不等式制約にて $g_i(x^*) \leq 0$, $i = 1, 2, \dots, \ell$, 等式制約にて $h_j(x^*) = 0$, $j = 1, 2, \dots, m$ を満たす点が線形独立制約想定を満たしているとは不等式制約の中で有効な制約式の勾配ベクトル $\nabla g_i(x^*)$, $i \in I(x^*)$ と等式制約の勾配ベクトル $\nabla h_j(x^*)$, $j = 1, 2, \dots, m$ が線形独立であるときをいう.

定理 7 (Karush-Kuhn-Tucker の最適性必要条件)

x^* が非線形計画問題 (4.2) の局所的最適解だとする. さらに x^* が線形独立制約想定を満たしていると仮定する. すると Karush-Kuhn-Tucker の最適性条件を満たす $(y^*, w^*) \in \mathbb{R}^\ell \times \mathbb{R}^m$ が存在する.

証明

[矢部 2006] などを参照.

□

定理 8 (2次の最適性必要条件)

x^* が非線形計画問題 (4.2) の局所的最適解だとする．さらに x^* が線形独立制約想定を満たしていると仮定する．すると Karush-Kuhn-Tucker の最適性条件を満たす $(y^*, w^*) \in \mathbb{R}^\ell \times \mathbb{R}^m$ が存在し，

$$u^T \left(\nabla^2 f(x^*) + \sum_{i \in I(x^*)} y_i^* \nabla^2 g_i(x^*) + \sum_{j=1}^m w_j^* \nabla^2 h_j(x^*) \right) u > 0,$$

が次の条件を満たす $u \neq 0$ に対していえる

$$\nabla g_i(x^*)^T u = 0, \quad i \in I(x^*), \quad \nabla h_j(x^*)^T u = 0, \quad j = 1, 2, \dots, m.$$

証明

[BERTSEKAS2003] など参照．

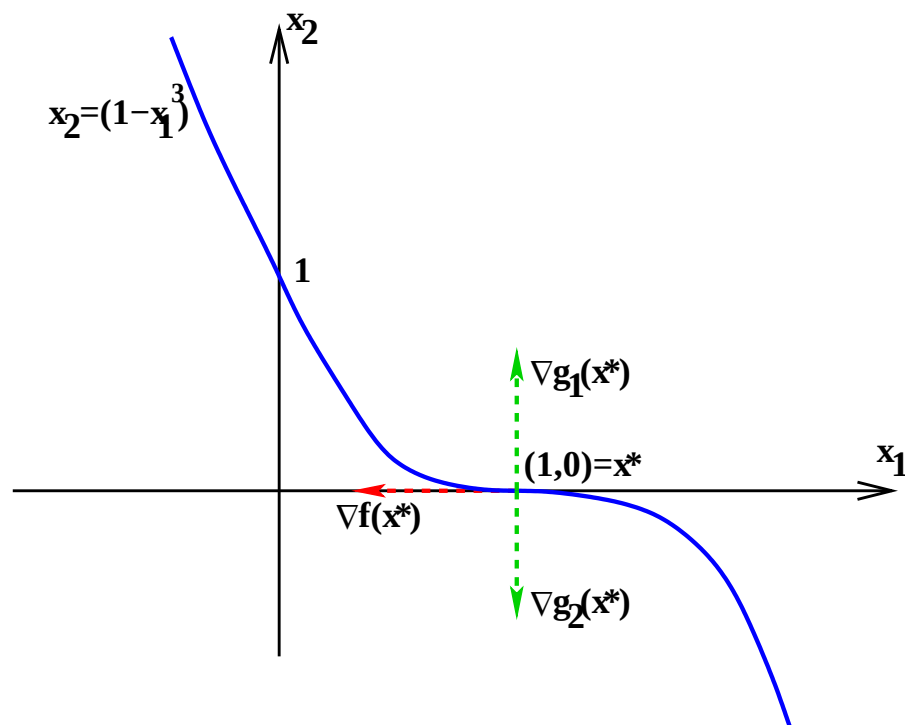
□

例 4

次の最適化問題を考える

$$\begin{cases} \text{最小化: } -x_1 \\ \text{条件: } -(1-x_1)^3 + x_2 \leq 0 \\ \quad \quad -x_2 \leq 0 \end{cases}$$

最適解は $(1, 0)$ であり，最適値は -1 である．



- 制約想定というのは数多くある.

定義 9

非線形計画問題 (4.2) において等式制約に相当する $h_j(x)$, $j = 1, 2, \dots, m$ が全て線形関数であり, 不等式制約に相当する $g_i(x)$, $i = 1, 2, \dots, \ell$ が全て凹関数である場合, この非線形計画問題は線形制約・凹制約想定を満たしているという.

定義 10

非線形計画問題 (4.2) の実行可能解 x^* が **Mangasarian-Fromovitz の制約想定** を満たしているとは等式制約の勾配ベクトル $\nabla h_j(x^*)$, $j = 1, 2, \dots, m$ が線形独立であり次の等式, 不等式を満たすベクトル $d \in \mathbb{R}^n$ が存在する時である.

$$(\nabla g_i(x^*))^T d < 0, \quad i \in I(x^*) \quad (\nabla h_j(x^*))^T d = 0, \quad j = 1, 2, \dots, m$$

例 5

$$\left\{ \begin{array}{l} \text{最小化: } f(x) \\ \text{条件: } \quad 2x_1 - 3x_2 + 1 = 0 \\ \quad \quad x_1 - 2x_2 + 1 = 0 \\ \quad \quad x_1^2 + x_2^2 - x_3 \leq 0 \end{array} \right.$$

$x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ において $d = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ が存在するので x^* は Mangasarian-Fromovitz の制約想定を満たす.

定義 1 1

非線形計画問題 (4.2) の実行可能解 x^* が **Slater の制約想定** を満たしているとは等式制約の関数 $h_j(x)$, $j = 1, 2, \dots, m$ が全て線形であり, 不等式制約の関数 $g_i(x)$, $i = 1, 2, \dots, \ell$ が全て凸関数でありかつ $g_i(\bar{x}) < 0$, $i \in I(x^*)$ を満たす $\bar{x} \in \mathbb{R}^n$ が存在した時である.

定理 1 2 (Karush-Kuhn-Tucker の最適性必要条件)

x^* が非線形計画問題 (4.2) の局所的最適解だとする. さらに x^* が線形制約・凹制約想定, Mangasarian-Fromovitz の制約想定, Slater の制約想定 of the いくつかを満たすとする. Karush-Kuhn-Tucker の最適性条件を満たす $(y^*, w^*) \in \mathbb{R}^\ell \times \mathbb{R}^m$ が存在する.

証明

[BERTSEKAS2003] などを参照.

□

(制約付き) 非線形計画問題

$$(4.2) \begin{cases} \text{最小化: } f(x) \\ \text{条件: } g_i(x) \leq 0, \quad i = 1, 2, \dots, \ell \\ \quad \quad h_j(x) = 0, \quad j = 1, 2, \dots, m \end{cases}$$

x^* が (4.2) の局所的最適解である

制約想定 必要条件 $\Downarrow$ $\Uparrow$ 十分条件 関数の凸性

次の条件を満たす (y^*, w^*) が存在する

Karush-Kuhn-Tucker の最適性条件

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{\ell} y_i^* \nabla g_i(x^*) + \sum_{j=1}^m w_j^* \nabla h_j(x^*) &= 0 \\ g_i(x^*) &\leq 0, \quad i = 1, 2, \dots, \ell \\ h_j(x^*) &= 0, \quad j = 1, 2, \dots, m \\ y_i^* g_i(x^*) &= 0, \quad i = 1, 2, \dots, \ell \\ y^* &\geq 0 \end{aligned}$$

4.4 数値解法

- 数値解法では 局所的最適解 を求めることが目的
- 大域的最適解を求めることは一般的に難しい

数値解法例

- 拡張 Lagrange 法
- 逐次2次計画法
- 信頼領域法
- 主双対内点法
- フィルター法

4.4.1 拡張Lagrange法

- まずは等式制約付き非線形計画問題について考える

$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } h_j(x) = 0, j = 1, 2, \dots, m \end{cases}$$

$\mu > 0$ をパラメータとした拡張Lagrange関数を定義する

$$\hat{\mathcal{L}}(x, w, \mu) \equiv f(x) + \sum_{j=1}^m w_j h_j(x) + \frac{\mu}{2} \sum_{j=1}^m h_j^2(x) \quad (4.3)$$

- $\mu \longrightarrow \infty$ によって等式制約がより強く課せられる
- 無制約非線形計画問題として考え, x に対して微分をし, 最適性条件を考える

$$\nabla_x \hat{\mathcal{L}}(x, w, \mu) = \nabla f(x) + \sum_{j=1}^m (w_j + \mu h_j(x)) \nabla h_j(x) = 0$$

さらに等式制約付き非線形計画問題に対するKKTの最適性条件(4.1)と比較すると

$$w_j^* \approx w_j + \mu h_j(x)$$

となることが推測できる. よって, この方法を反復的に行なって(4.3)を満たすような (x^*, w^*) を求めようとする,

$$w_j^{k+1} := w_j^k + \mu^k h_j(x^k)$$

と Lagrange 乗数を更新することが自然である.

拡張 Lagrange 法

Step 0: 初期点 x^0 , w^0 を決め, 十分大きい $\mu^0 > 0$, $k := 0$ とする.

終了条件: もし 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 (w^k とともに KKT の最適性条件を満たしているならば) 終了する.

Step 1: (無制約) 非線形計画問題に対する手法を用いて $\hat{\mathcal{L}}(x, w^k, \mu^k)$ を x^k から出発して最小化することを試みて x^{k+1} を求める. ただし, w^k, μ^k は固定定数とみなす.

Step 2: Lagrange 乗数を $w_j^{k+1} := w_j^k + \mu^k h_j(x^{k+1})$ として更新する.

Step 3: $\mu^{k+1} > \mu^k$ を決め, $k := k + 1$ とし, 終了条件を調べる.

定理 1 3

等式制約付き非線形計画問題の局所的最適解 x^* とそれに対応する Lagrange 乗数 w^* が 2 次の最適性十分条件を満たしていると仮定する (定理 5 参照). さらに局所的最適解 x^* に対する線形独立制約想定が成り立っていると仮定する. この時, 十分大きな $\mu^* > 0$ が存在して, $\mu > \mu^*$ に対して $\nabla_x^2 \hat{\mathcal{L}}(x^*, w^*, \mu)$ は正定値行列になる.

証明

[矢部 2006] などを参照.

□

- 不等式制約付き非線形計画問題を考える

$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } g_i(x) \leq 0, \quad i = 1, 2, \dots, \ell \end{cases}$$

- この場合, **スラック変数** $\mathbb{R}^\ell \ni s > 0$ を導入して等式制約付き非線形計画問題に変換する

$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } g_i(x) + s_i = 0, \quad i = 1, 2, \dots, \ell \\ \quad \quad s_i \geq 0, \quad \quad \quad i = 1, 2, \dots, \ell \end{cases}$$

- 等式制約問題なので，もとの等式制約の記述を用い，正確には箱型制約と等式制約付き非線形計画問題を扱う．ただし， $\mathbb{R}^n \ni b \leq u \in \mathbb{R}^n$ は変数 x が取り得る値の範囲である

$$\left\{ \begin{array}{l} \text{最小化: } f(x) \\ \text{条件: } h_j(x) = 0, \quad j = 1, 2, \dots, m \\ b \leq x \leq u \end{array} \right.$$

上記の問題に対する拡張 Lagrange 関数は

$$\hat{\mathcal{L}}(x, w, \mu) = f(x) + \sum_{j=1}^m w_j h_j(x) + \frac{\mu}{2} \sum_{j=1}^m h_j^2(x)$$

であり，拡張 Lagrange 法の部分問題では次の最適化問題を考える

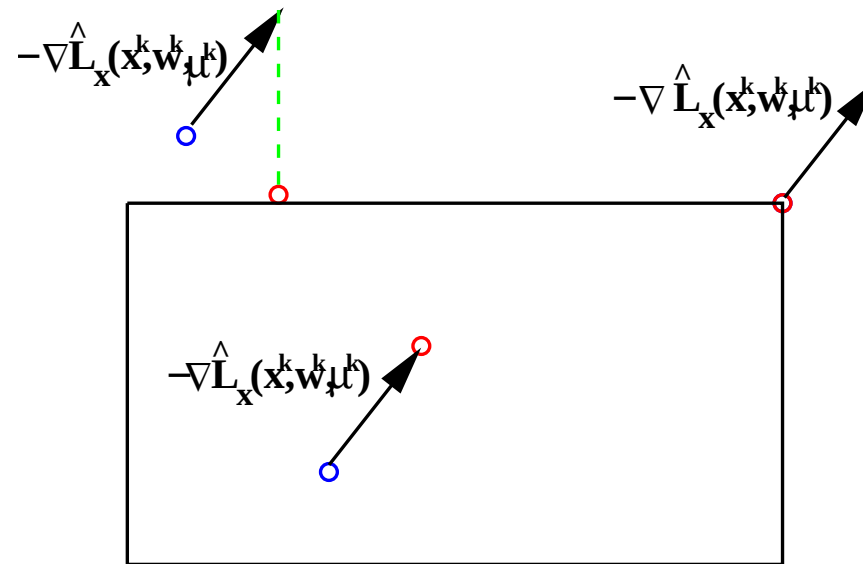
$$\left\{ \begin{array}{l} \text{最小化: } \hat{\mathcal{L}}(x, w, \mu) = f(x) + \sum_{j=1}^m w_j h_j(x) + \frac{\mu}{2} \sum_{j=1}^m h_j^2(x) \\ \text{条件: } b \leq x \leq u \end{array} \right.$$

上記の問題は例えば勾配ベクトル射影法で解くことができる．つまり次のような関係式を満たす x^k を求めればよい．

$$x^k := \mathcal{P}(x^k - \nabla \hat{\mathcal{L}}_x(x^k, w^k, \mu^k); b, u)$$

ただし, $\mathcal{P}(v; b, u)$ はベクトル v を箱型制約に射影した点である.

$$\mathcal{P}(v; b, u) = \begin{cases} b_i, & \text{if } v_i \leq b_i, \\ v_i, & \text{if } b_i \leq v_i \leq u_i, \quad i = 1, 2, \dots, n \\ u_i, & \text{if } u_i \leq v_i \end{cases}$$



よって上記のルーチンを拡張Lagrange法アルゴリズムのStep 1に組み込むことによって不等式制約付き非線形計画問題も解けるようになる.

4.4.2 逐次2次計画法

- まずは等式制約付き非線形計画問題について考える

$$\begin{cases} \text{最小化: } f(x) \\ \text{条件: } h_j(x) = 0, j = 1, 2, \dots, m \end{cases}$$

上記の最適化問題に対する KKT の最適性条件 (4.1) から次の方程式に対する Newton 法を考える

$$F(x, w) = \begin{pmatrix} \nabla f(x) + \sum_{j=1}^m w_j \nabla h_j(x) \\ h(x) \end{pmatrix} = 0$$

つまり, (x^k, w^k) を k 反復目の点だとしたら, 次点は次の式で求まる

$$\begin{pmatrix} x^{k+1} \\ w^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ w^k \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix}, \quad \text{ただし} \quad \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix} = -[F'(x^k, w^k)]^{-1} F(x^k, w^k),$$

$$F'(x^k, w^k) = \begin{pmatrix} \nabla^2 f(x^k) + \sum_{j=1}^m w_j^k \nabla h_j^2(x^k) & \nabla h(x^k) \\ \nabla h^T(x^k) & O \end{pmatrix}$$

なので, 上記の行列の逆行列は次の仮定が満たされていれば存在する [NOCEDAL2006].

仮定

- (1) $\nabla h_j(x^k)$, $j = 1, 2, \dots, m$ は線形独立である.
- (2) $u^T \left(\nabla^2 f(x^k) + \sum_{j=1}^m w_j^k \nabla^2 h_j(x^k) \right) u > 0$, が次の条件を満たす全ての $u \neq 0$ に対していえる $\nabla h_j(x^k)^T u = 0$, $j = 1, 2, \dots, m$.

前の式を展開する以下の式が得られる.

$$\begin{cases} \left[\nabla^2 f(x^k) + \sum_{j=1}^m w_j^k \nabla^2 h_j(x^k) \right] \Delta x + \sum_{j=1}^m (\Delta w)_j \nabla h_j(x^k) = -\nabla f(x^k) - \sum_{j=1}^m w_j^k \nabla h_j(x^k) \\ \nabla h_j^T(x^k) \Delta x = -h_j(x^k), \quad j = 1, 2, \dots, m \end{cases}$$

逐次2次計画法

Step 0: 初期点 x^0 , w^0 を決め, $k = 0$ とする.

終了条件: もし 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 (w^k とともに KKT の最適性条件を満たしているならば) 終了する.

Step 1: 上記の式より $(\Delta x, \Delta w)$ を求める.

Step 2: $(x^{k+1}, w^{k+1}) = (x^k, w^k) + (\Delta x, \Delta w)$, $k = k + 1$ とし, 終了条件を調べる.

- 探索方向 $(\Delta x, \Delta w)$ はまた下の問題の目的関数と制約式の近似モデルに対する KKT の最適化条件としても見なすことができる

$$\begin{cases} \text{最小化: } \mathcal{L}(x, w^k) = f(x) + \sum_{j=1}^m w_j^k h_j(x) \\ \text{条件: } h_j(x) = 0, & j = 1, 2, \dots, m \end{cases}$$

↓ x^k における近似

$$\begin{cases} \text{最小化: } \nabla_x \mathcal{L}(x^k, w^k) \Delta x + \frac{1}{2} \Delta x \nabla_x^2 \mathcal{L}(x^k, w^k) \Delta x \\ \text{条件: } h_j(x^k) + \nabla h_j^T(x^k) \Delta x = 0, & j = 1, 2, \dots, m \end{cases}$$

↓ Δx に対する KKT の最適性条件

$$\begin{cases} \nabla_x^2 \mathcal{L}(x^k, w^k) \Delta x + \nabla_x \mathcal{L}(x^k, w^k) + \sum_{j=1}^m (\Delta w)_j \nabla h_j(x^k) = 0 \\ h_j(x^k) + \nabla h_j^T(x^k) \Delta x = 0, & j = 1, 2, \dots, m \end{cases}$$

- 一般の非線形計画問題を考える

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

- 同様に目的関数と制約式の近似モデルを考える

$$\begin{cases} \text{最小化: } \nabla_x \mathcal{L}(x^k, w^k) \Delta x + \frac{1}{2} \Delta x \nabla_x^2 \mathcal{L}(x^k, w^k) \Delta x \\ \text{条件: } g_i(x^k) + \nabla g_i^T(x^k) \Delta x \leq 0, & i = 1, 2, \dots, \ell \\ h_j(x^k) + \nabla h_j^T(x^k) \Delta x = 0, & j = 1, 2, \dots, m \end{cases}$$

- もし上記の問題の $\nabla_x^2 \mathcal{L}(x^k, w^k)$ が半正定値行列になると、この最適化問題は変数 Δx に対する凸 2 次計画問題になるので、単体法や内点法で簡単に解くことができる
- 以上のルーチンを逐次 2 次計画法の **Step 1** に組み込むことによって一般の非線形計画問題に対しても対応できる
- $\nabla_x^2 \mathcal{L}(x^k, w^k)$ の代わりにその近似行列 B^k を用いることもできる