4.6　不等式制約問題

今まで制約式が等式の場合を扱ってきたが，より一般的に等式と不等式を制約に持つ場合を考える．

例 **22** (射影問題). 平面 $4x + y + 2z = 2$ と単位球の内部 $x^2 + y^2 + z^2 \leq 1$ との共通部分の点で，点 $(2, 3, 4)$ までの距離が一番近い点を求めよ．

この問題は

$$\begin{aligned} \text{最小化}\ & f(x,y,z) := (x-2)^2 + (y-3)^2 + (z-4)^2 \\ \text{制約}\ & g_1(x,y,z) := x^2 + y^2 + z^2 - 1 \leq 0 \\ & g_2(x,y,z) := 4x + y + z - 2 = 0 \end{aligned}$$

と定式化できる．すると制約式に不等式と等式が現れる．

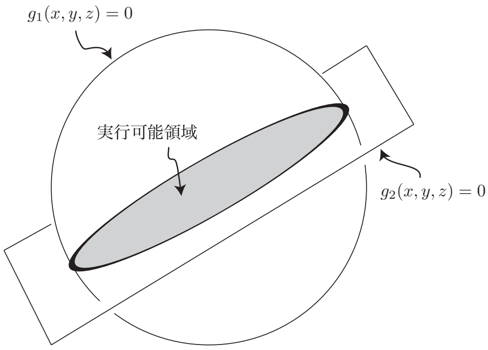


図 4.12: 不等式と等式で決まる領域（斜線の部分）

4.6.1　不等式が一つの場合

この問題の最適性条件について，まず制約式が円周とその内部を表す不等式一つのみの場合を解説する．

$$\begin{aligned} (P) \qquad \text{最小化}\ & f(x,y) \\ \text{制約}\ & g(x,y) := x^2 + y^2 - 1 \leq 0 \end{aligned}$$

(a, b) を上の最適化問題の局所最小解とすると，次の二つの場合が考えられる；

(1) (a, b) が円の内部にある $(g(a, b) < 0)$ 場合．このとき

$$\nabla f(a, b) = \mathbf{0}$$

が成り立つ（通常の停留点）．

解説．(a,b) が制約無しの最小化問題

$$(P')\quad \text{最小化}\ f(x,y) \\ \text{制約なし}$$

の局所最小解になっていることを示そう．　まず，次の例を見てみよう．

$$\begin{cases} \text{最小化} & x^2+y^2 \\ \text{制約} & x^2+y^2-1 \le 0 \end{cases} \qquad \begin{cases} \text{最小化} & x^2+y^2 \\ \text{制約} & \text{なし} \end{cases}$$

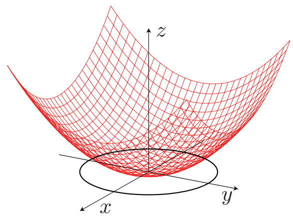


図 4.13: $z = x^2+y^2$ のグラフ

いま左側の問題の最小解は $(0,0)$ であるが，最小解が実行可能解の内部に含まれているので，右側の問題のように制約を外しても，最小解は同じく $(0,0)$ になる．

この例を参考にして，問題 (P) について考えよう．いま，(a,b) の十分近くにある任意の (x,y) を考える．ここで十分近くとは，$g(x,y)<0$ を満たした上で十分近いとする（図 4.14 左の一番小さい等高線の輪くらい）．すると (x,y) は実行可能領域の内部にあるので，(a,b) が (P) の局所最小解であるという仮定より，

$$f(x,y) \ge f(a,b) \quad ((a,b)\ \text{に十分近いすべての}\ (x,y))$$

が成り立つ．したがって，(a,b) は (P') の局所最小解になっている．よって $\nabla f(a,b) = \mathbf{0}$ が成り立つ．

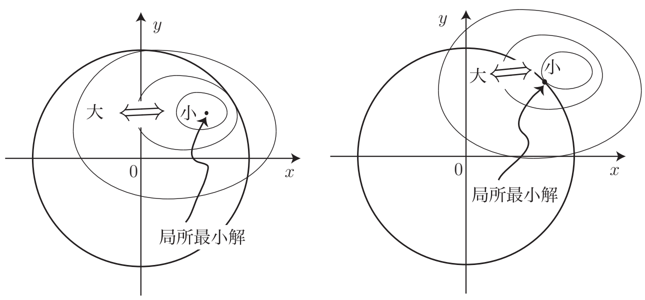


図 4.14: 実行可能領域（円の周上とその内部）目的関数の等高線の関係：左図は (P) の局所最小解が円の内部，右図は円周上にある場合

(2) <u>(a,b) が円周上にある $(g(a,b)=0)$ 場合</u>．このとき，ある $\lambda \ge 0$ が存在して，

$$-\nabla f(a,b) = \lambda \nabla g(a,b)$$

が成り立つ（λ の符号が ≥ 0 であることに注意）．

解説．まず λ の符号を考えずに，ある $\lambda \in \mathbb{R}$ が存在して，$f(a,b) = \lambda \nabla g(a,b)$ が成り立つことを示そう．いま，特に (a,b) に近い円周上の (x,y) に対しても，$f(x,y) \ge$

$f(a,b)$ が成り立つ．したがって，(a,b) は

$$(P'') \quad \text{最小化}\ f(x,y) \quad \text{制約}\ x^2+y^2=1$$

の局所最適解にもなる．よって等式制約問題の最適性条件（定理 9）より，ある $\lambda \in \mathbb{R}$ が存在して，$\nabla f(a,b) = \lambda \nabla g(a,b)$ が成り立つ．

次に，λ の符号が ≥ 0 であることを示す．このために，点 (a,b) を始点としたとき勾配ベクトル $\nabla g(a,b)$ がどちらの方向を向いているか調べよう．いま，制約式は

$$g(x,y) \leq 0$$

なので，実行可能領域の境界は $g(x,y)=0$ となる等高線で，実行可能領域の内部は g の値が小さく，外部は g の値が大きい領域になっている．ここで，定理 9 の直後

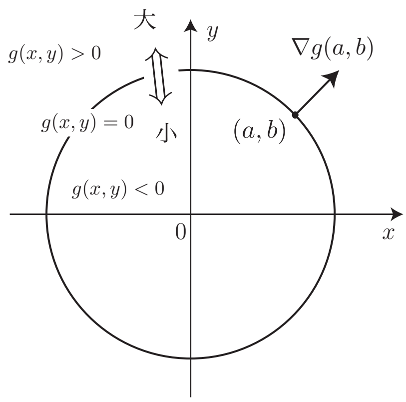


図 4.15: 勾配ベクトルの向きと不等式領域

の解説より，一般的に $\nabla g(a,b)$ は g の等高線に直交し，g の値が増える方向を向いていることがわかっている．したがって, 点 (a,b) において，$\nabla g(a,b)$ は実行可能領域の外側を向き，境界に直交した方向を向いている．

一方 $\{-\nabla f(a,b)\}$ は目的関数 f の等高線に直交し値が減る方向なので，図 4.14 の右図より，それも実行可能領域の外側を向いていることになる．従って，二つのベクトルが同じ方向を向いていることから

$$-\nabla f(a,b) = \lambda \nabla g(a,b),\ \lambda \geq 0$$

となる．

上記二つの場合をまとめて書くと，

定理 **12.** 最小化問題

$$\text{最小化}\ f(x) \quad \text{制約}\ g(x) \leq 0$$

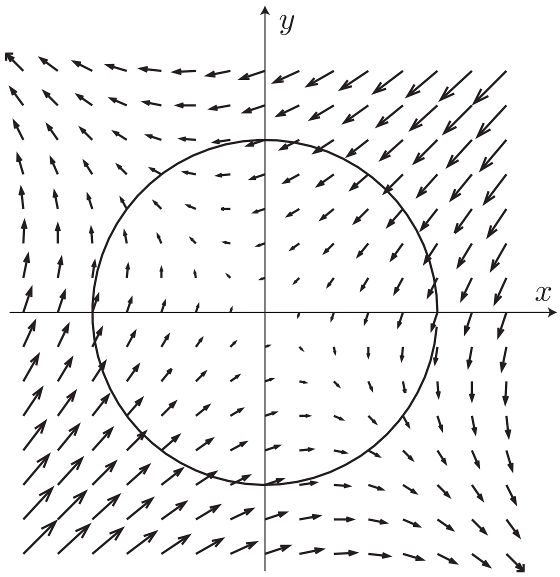


図 4.16: $f(x,y) = x^2 + 6xy + y^2$ としたとき，各点でのベクトル $(-1)\times\ \nabla f(x,y)$

に対して，$\bar{x}$ が局所最小解であり，$\nabla g(\bar{x}) \neq \mathbf{0}$ ならば，ある数 λ が存在して，

$$(*)\quad \begin{cases} -\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}) \\ \lambda g(\bar{x}) = 0, \lambda \geq 0 \\ g(\bar{x}) \leq 0 \end{cases}$$

が成り立つ．

解説. 具体的な制約式について解説したが一般の制約式に対しても，同様のことが成り立つ．$\bar{x}$ を局所最小解とする．$g(\bar{x}) < 0$ のときは，上記議論の (1) より $\nabla f(\bar{x}) = 0$ となるので，$\lambda = 0$ とおけば $(*)$ を満たす．$g(\bar{x}) = 0$ のときは上記議論の (2) より，ある $\lambda \geq 0$ が存在して，$-\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$ となるので，やはり $(*)$ を満たす．

4.6.2　不等式が複数あるとき

$$\begin{aligned} (P) \qquad \text{最小化}\ & f(x,y) := x^2 + 6xy + y^2 \\ \text{制約}\ & g_1(x,y) := x^2 + y^2 - 1 \leq 0 \\ & g_2(x,y) := -x - 1/2 \leq 0 \end{aligned}$$

で，(a,b) を局所最小解とする．(a,b) の位置について，以下三つの場合がある．

(1). (a,b) が $g_1(a,b) < 0$，$g_2(a,b) < 0$ をみたす場合，

$$\nabla f(a,b) = \mathbf{0}$$

となる．

解説．このとき，(a,b) は実行可能領域の内部にある．不等式が一つの場合と同様の議論により，制約無しの最小問題

$$\text{最小化 } f(x,y)$$

の局所最小解になっていることから導かれる．

(2). (a,b) が $g_1(a,b) = 0$，$g_2(a,b) < 0$ をみたす場合．

$$-\nabla f(a,b) = \lambda_1 \nabla g_1(a,b) \text{ かつ } \lambda_1 \geq 0$$

となる．

解説．このとき，(a,b) は

$$(P'') \quad \begin{aligned} &\text{最小化 } f(x,y) \\ &\text{制約 } g_1(x,y) \leq 0 \end{aligned}$$

の局所最小解になることから導かれる．

(3). (a,b) が $g_1(a,b) < 0$，$g_2(a,b) = 0$ をみたす場合，

$$-\nabla f(a,b) = \lambda_2 \nabla g_2(a,b) \text{ かつ } \lambda_2 \geq 0$$

となる．

解説．(2) と同様．

(4). (a,b) が $g_1(a,b) = 0$，$g_2(a,b) = 0$ をみたす場合，

$$-\nabla f(a,b) = \lambda_1 \nabla g_1(a,b) + \lambda_2 \nabla g_2(a,b) \text{ かつ } \lambda_1 \geq 0,\ \lambda_2 \geq 0$$

となる．

解説．点 (a,b) は，それぞれ $g_1(x,y) = 0$ と $g_2(x,y) = 0$ で与えられる曲線の交点に位置する．まず，$\nabla g_1(a,b)$ と $\nabla g_2(a,b)$ は，不等式一つの場合と同様に考えると実行可能領域の外側を向いている．いま，(a,b) は局所最小解なので，目的関数 f の等高線はそこで実行可能領域の境界に接している．また，$-\nabla f(a,b)$ はその等高線と直交し，値が減る方向を向いている．これは，図 4.17 と図 4.18 より，$-\nabla f(a,b)$ が二つのベクトル $\nabla g_1(a,b)$ と $\nabla g_2(a,b)$ に挟まれた方向にあるということである．これを式で書くと

$$-\nabla f(a,b) = \lambda_1 \nabla g_1(a,b) + \lambda_2 \nabla g_2(a,b) \quad \text{かつ} \quad \lambda_1 \geq 0,\ \lambda_2 \geq 0$$

となる．

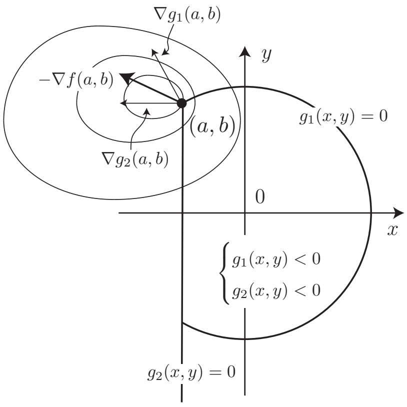


図 4.17: 目的関数の勾配ベクトルが制約式の勾配ベクトルに挟まれている

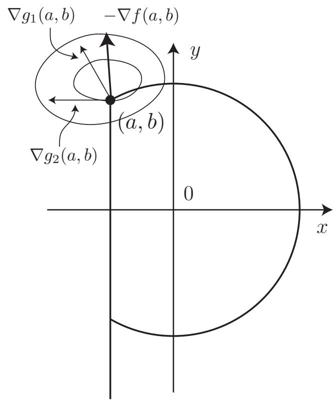


図 4.18: 制約式の勾配ベクトルに挟まれていない．このとき目的関数の等高線と実行可能領域の境界は交わる．

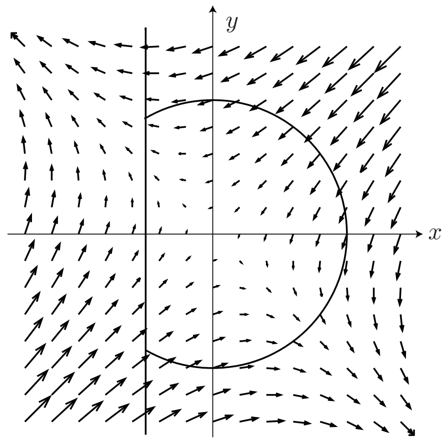


図 4.19: 勾配ベクトル $-\nabla f(x,y)$ と制約 $x^2+y^2-1\leq 0, -x-1/2\leq 0$

三つの場合をまとめて書くと，

定理 13. 最適化問題

$$\begin{aligned} \text{最小化}\ & f(\bar{x}) \\ \text{制約}\ & g_1(\bar{x})\leq 0,\ g_2(\bar{x})\leq 0 \end{aligned}$$

に対して，$\bar{x}$ が局所最小解であり，$\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x})\}$ が一次独立であるとする．すると，ある数 λ_1，λ_2 が存在して，

$$\begin{cases} -\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0,\ \lambda_i \geq 0 \\ g_i(\bar{x}) \leq 0\ (i=1,2) \end{cases}$$

が成り立つ.

4.6.3　不等式と等式があるとき

制約式に不等式と等式がある場合は次のようになる.

定理 14. 最適化問題

$$
\begin{aligned}
\text{最小化}\ & f(x) \\
\text{制約}\ & g_1(x) \leq 0, g_2(x) \leq 0 \\
& h(x) = 0
\end{aligned}
$$

に対して, $\bar{x}$ が局所最小解であり $\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x}), \nabla h(\bar{x})\}$ が一次独立であるとする．すると, ある数 λ_1, λ_2, μ が存在して,

$$
(*)\quad \begin{cases} -\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) + \mu \nabla h(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0,\ \lambda_i \geq 0,\ g_i(\bar{x}) \leq 0 \\ \mu : \text{任意},\ h(\bar{x}) = 0 \end{cases}
$$

が成り立つ. この条件式を KKT 条件 と呼ぶ.

KKT 条件の解説

具体例を用いて解説しよう. α, β, γ, δ を定数とする.

$$
\text{最小化}\ f(x,y,z) = \alpha x + \beta y + \gamma z + \delta
$$

$$
\text{制約}\ \begin{cases} g_1(x,y,z) = x^2 + y^2 + z^2 - 1 \leq 0 \\ g_2(x,y,z) = -x - 1/2 \leq 0 \\ h(x,y,z) = x + y + z - 1 = 0 \end{cases}
$$

という, 一次関数を最小化する問題を考える. 実行可能領域は, 空間内の円を図 4.20 のように切り取った部分になる. いままでと同様に, 点 $\bar{x} = (a,b,c)$ が局所最小解ならば, 目的関数の等高面と実行可能領域が接する.

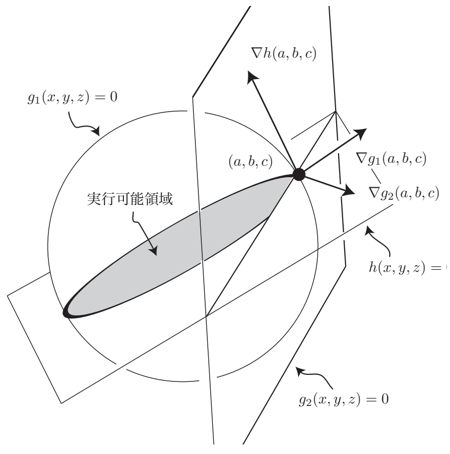


図 4.20: 実行可能領域（空間内の円）

点 (a,b,c) が空間内の円の内側の方にある場合

まず，この場合 $(g_1(a,b,c) < 0, g_2(a,b,c) < 0, h(a,b,c) = 0)$ は 4.6.1 節の (1) と同様に考えると，点 (a,b,c) は

$$\begin{aligned} \text{最小化 } & f(x,y,z) \\ \text{制約 } & h(x,y,z) = 0 \end{aligned}$$

の局所最小解になる．したがって，ある μ が存在して，

$$\nabla f(a,b,c) = \mu \nabla h(a,b,c)$$

が成り立つ．よって，$\lambda_1 = \lambda_2 = 0$ とおけば $(*)$ を満たす．

点 (a,b,c) が空間内の円の境界にある場合

次に，この場合 $(g_1(a,b,c) = 0, g_2(a,b,c) < 0, g_3(a,b,c) = 0)$ は 4.6.1 節の (1) と同様に考えると，点 (a,b,c) は

$$\begin{aligned} \text{最小化 } & f(x,y,z) \\ \text{制約 } & \begin{cases} g_1(x,y,z) = 0 \\ h(x,y,z) = 0 \end{cases} \end{aligned}$$

の局所最小解になる．したがって，定理 **??** より

$$-\nabla f(a,b,c) = \lambda_1 \nabla g_1(x,y,z) + \mu \nabla h(a,b,c)$$

となる．さらに，4.6.1 節の (2) の議論より，$\lambda_1 \geq 0$ と言える．これより，$\lambda_2 = 0$ とおけば，$(*)$ を満たす．

点 (a,b,c) が実行可能領域の角張ったところにある場合

最後に，点 (a,b,c) が角張ったところにある場合（図4.21）を考える．いま，点 (a,b,c) で実行可能領域に接する平面は図 4.21 のように回転させても実行可能領域に接し続ける（横の回転には制限あり）．よって，実行可能領域に接する平面の法線ベクトルは，回転させた平面のものも含めて，

$$\lambda_1 \nabla g_1(a,b,c) + \lambda_2 \nabla g_2(a,b,c) + \mu \nabla h(a,b,c), \\ \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \mu : \text{任意}$$

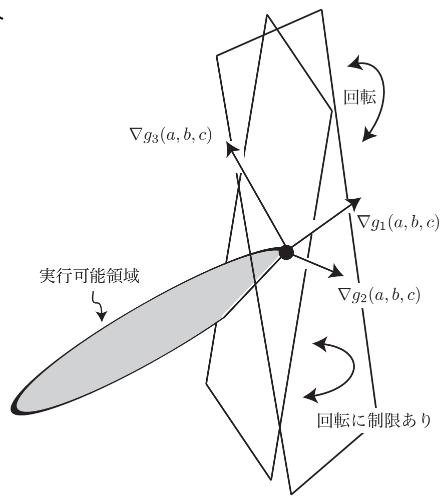


図 4.21: 実行可能領域と接する平面

となる．

いま，目的関数の (a,b,c) を通る等高面の法線ベクトルは $\nabla f(a,b,c)$ なので，

$$-\nabla f(a,b,c) = \lambda_1 \nabla g_1(a,b,c) + \lambda_2 \nabla g_2(a,b,c) + \mu \nabla h(a,b,c), \qquad \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \mu : \text{任意}$$

と書ける．

まとめと補足

以上の場合をまとめると定理の式が成り立つ．KKT 条件の書き方として，次の書き方は良く出てくるのでここで紹介しておく．

補足. λ_i に関する条件を，ベクトル $\lambda = {}^t(\lambda_1, \lambda_2)$，$g(x) = {}^t(g_1(x), g_2(x))$ と内積を用いると，

$$\begin{cases} \lambda \cdot g(\bar{x}) = 0, \\ \lambda_1 \geq 0,\ g_1(\bar{x}) \leq 0, \\ \lambda_2 \geq 0,\ g_2(\bar{x}) \leq 0 \end{cases}$$

と書ける．これを特に相補性条件 と呼ぶ．

もっと制約が多い場合

制約式の数が多い場合も同様の定理が成り立つ．

定理 15.

$$\begin{aligned} \text{最小化}\ & f(x) \\ \text{制約}\ & g_1(x) \leq 0, \ldots, g_m(x) \leq 0 \\ & h_1(x) = 0, \ldots, h_l(x) = 0 \end{aligned}$$

に対して，$\bar{x}$ が局所最小解であり $\{\nabla g_1(\bar{x}), \ldots, \nabla g_m(\bar{x}), \nabla h_1(\bar{x}), \ldots, \nabla h_l(\bar{x})\}$ が一次独立であるとする ．すると，ある数 $\lambda_1, \ldots, \lambda_m$，$\mu_1, \ldots, \mu_l$ が存在して，

$$(*) \qquad \begin{cases} -\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \ldots + \lambda_m \nabla g_m(\bar{x}) \\ \qquad\qquad\qquad + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_l \nabla h_l(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0,\ \lambda_i \geq 0,\ g_i(\bar{x}) \leq 0 \\ \mu_j : \text{任意},\ h_j(\bar{x}) = 0 \end{cases}$$

が成り立つ．

4.6.4　例題

次の最小化問題を解く．

$$\begin{aligned} \text{最小化}\ & f(x,y) = x^2 + 6xy + y^2 \\ \text{制約}\ & g(x,y) = x^2 + y^2 - 1 \leq 0 \end{aligned}$$

まず KKT 条件を書くと

$$(*)\ \begin{cases} -\begin{bmatrix} 2x+6y \\ 6x+2y \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix} \\ \lambda(x^2+y^2-1) = 0,\ \lambda \geq 0 \\ x^2+y^2-1 \leq 0 \end{cases}$$

となる．以下場合分けで解く．

$x^2 + y^2 - 1 < 0$ の場合，$(*)$ の第 3 式を満たすものは $\lambda = 0$，$(x,y) = (0,0)$ となる．これは $(*)$ を満たす．

次に $x^2 + y^2 - 1 = 0$ の場合を考える．まず

$$-\begin{bmatrix} 2x+6y \\ 6x+2y \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix},\ x^2+y^2-1=0$$

を満たす点を探すと，固有値問題を解いて

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \mp\frac{1}{\sqrt{2}} \\ 2 \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \mp\frac{1}{\sqrt{2}} \\ -4 \end{bmatrix}$$

を得る．さらに，この内 $\lambda \geq 0$ を満たすものは前者になる．

よって KKT 条件を見たす点は，$(x,y) = (0,0), (\pm 1/\sqrt{2}, \mp 1/\sqrt{2})$ となるので，それぞれにおける目的関数 f の値を調べると，最小値は

$$f\left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}\right) = -2$$

となる．

例 **23.** 最適化問題

$$\begin{aligned} \text{最小化}\ & f(x,y,z) := (x-2)^2 + (y-3)^2 + (z-4)^2 \\ \text{制約}\ & x^2+y^2+z^2 \leq 1 \\ & 4x + y + 2z = 2 \end{aligned}$$

の KKT 条件は

$$\begin{cases} -\begin{bmatrix} 2x-4 \\ 2y-6 \\ 2z-8 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} + \mu \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} \\ \lambda(x^2+y^2+z^2-1)=0, \lambda \geq 0, \mu : \text{任意} \\ x^2+y^2+z^2-1 \leq 0, 4x+y+2z-2=0 \end{cases}$$

となる.

練習問題 5. 1. 以下の最小化問題の最小値を求めよ

(1).

$$\begin{aligned} \text{最小化}\ & f(x,y) = x^2+xy+y^2 \\ \text{制約}\ & g(x,y) = x^2+y^2-1 \leq 0 \end{aligned}$$

(2).

$$\begin{aligned} \text{最小化}\ & f(x,y) = xy \\ \text{制約}\ & g(x,y) = x^2+y^2-1 \leq 0 \end{aligned}$$

2. 以下の最小化問題の KKT 条件を求めよ（解は求めなくてよい）.

$$\begin{aligned} \text{最小化}\ & x-y+2z \\ \text{制約}\ & x^2+y^2+2z^2-1 \leq 0 \\ & 2x+2y+z-1=0 \end{aligned}$$

3. 原点を中心とする半径 $\sqrt{3}$ の球と平面 $x+2y+2z=3$ の交点の中で，x 座標が最小となるものを求めよ.

4.7　凸計画

目的関数も制約式もすべて凸関数である問題を凸計画と呼ぶ．凸計画は有用な性質を持っている．

定理 16.

$$\begin{array}{lll} (P) & \text{最小化} & f(x) \\ & \text{制約} & g_i(x) \le 0,\ i = 1, \ldots, m \end{array}$$

で，f，g_i は凸関数であるとする．$\bar{x}$ が KKT 条件を満たせば $\bar{x}$ は (P) の大域最小解である．

証明．f は凸関数なので，

$$f(x) - f(\bar{x}) \ge \nabla f(\bar{x}) \cdot (x - \bar{x})$$

が成り立つ．また，$\bar{x}$ は KKT 条件を満たすので，ある $\lambda_i \ge 0$ が存在して

$$\nabla f(\bar{x}) = -\sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x})$$

が成り立ので，

$$f(x) - f(\bar{x}) \ge -\sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) \cdot (x - \bar{x})$$

を得る．ここで右辺が 0 以上になることを示す．まず，g_i も凸関数なので

$$g_i(x) - g_i(\bar{x}) \ge \nabla g_i(\bar{x}) \cdot (x - \bar{x})$$

が成り立つ．いま x を実行可能解とすると $g_i(x) \le 0$ となる．$g_i(\bar{x}) = 0$ のときは，

$$0 \ge g_i(x) = g_i(x) - g_i(\bar{x}) \ge \nabla g_i(\bar{x}) \cdot (x - \bar{x})$$

なので，$-\lambda_i \nabla g_i(\bar{x}) \cdot (x - \bar{x}) \ge 0$ となる．また，$g_i(\bar{x}) \le 0$ のときは，$\lambda_i = 0$ なので，$-\lambda_i \nabla g_i(\bar{x}) \cdot (x - \bar{x}) = 0$ となる．したがって，任意の実行可能界 x に対して，$f(x) - f(\bar{x}) \ge 0$ である．　□

まとめ

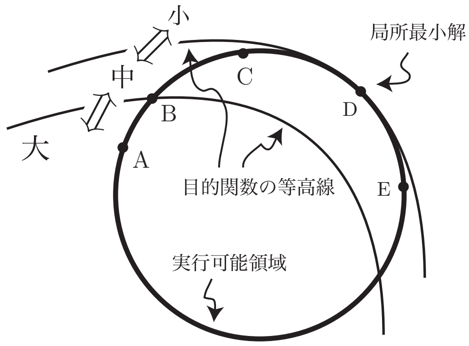
目的関数の等高線は最適解で実行可能領域に接する

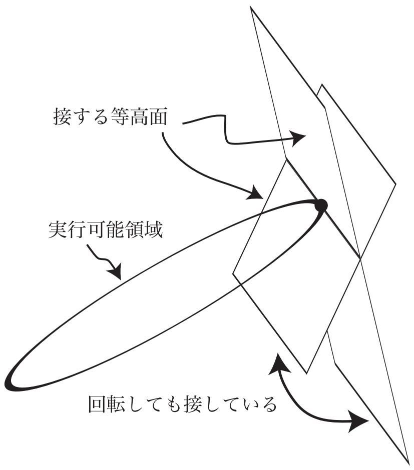
等式制約

最小化　$f(x, y)$

制約　$g(x, y) = 0$

$$\Longrightarrow \quad \nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$$





複数の等式制約

最小化　$f(x, y, z)$

制約　$g_1(x, y, z) = 0$

$g_2(x, y, z) = 0$

$$\Longrightarrow \quad \nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 g_2(\bar{x})$$

不等式制約

最小化　$f(x, y)$

制約　$g_1(x, y) \leq 0$

$g_2(x, y) \leq 0$

$$\Longrightarrow \begin{cases} -\nabla f(a, b) = \lambda_1 \nabla g_1(a, b) + \lambda_2 \nabla g_2(a, b) \\ \lambda_1 \geq 0, \lambda_2 \geq 0 \end{cases}$$