不等式制約の最適化

§ 等式制約の最適化

問題:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad g(\boldsymbol{x}) = 0 \qquad (1)$$

: 目的関数

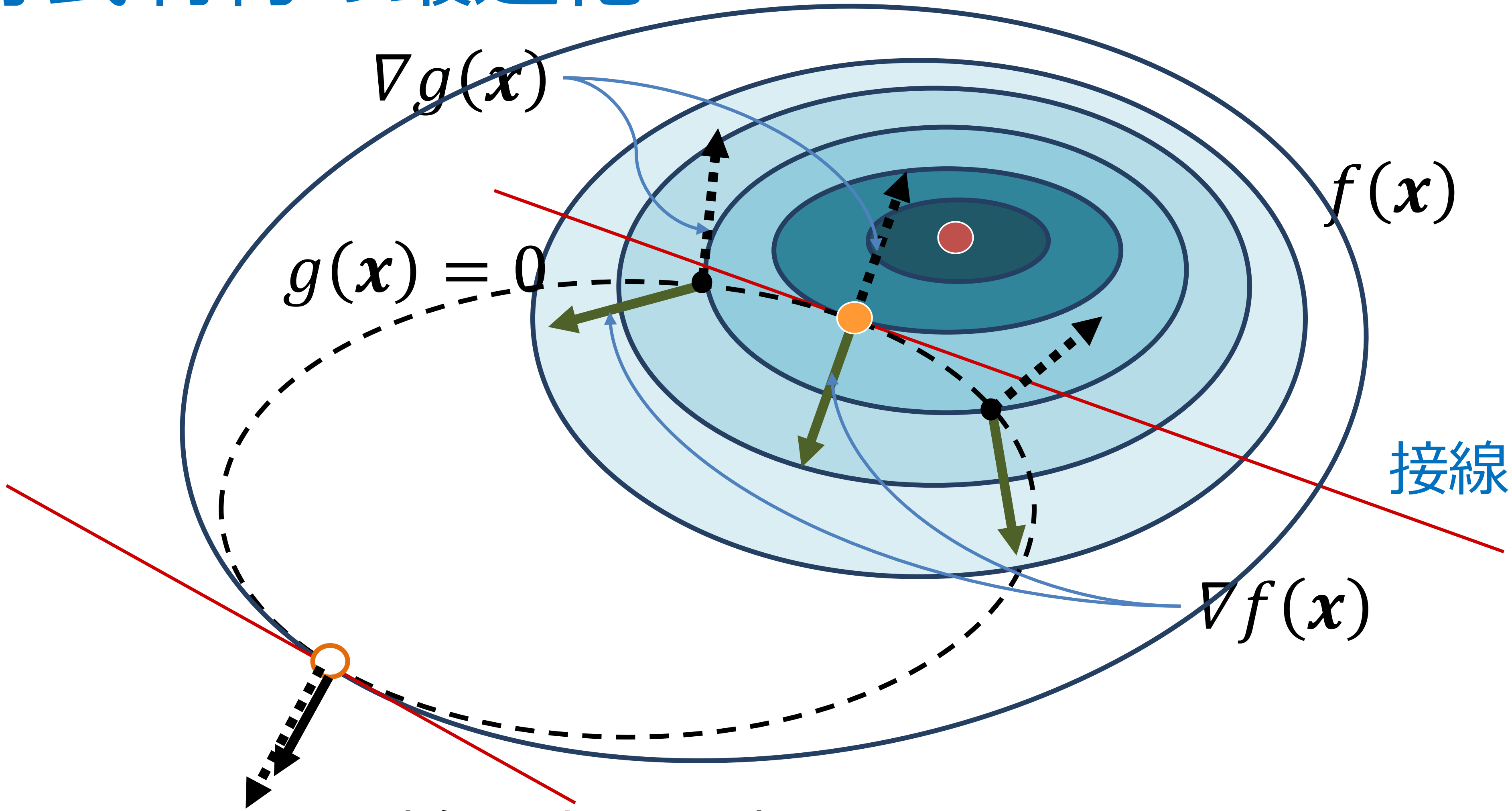
: 制約

ここで,

$\boldsymbol{x} \in R^n$: 設計変数

$f(\boldsymbol{x}), g(\boldsymbol{x})$ は, $R^n \to R$の関数

等式制約の最適化



最適解を与える点 ● において，

$\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0, g(\boldsymbol{x}) = 0.$ （必要条件）

§ 等式制約の最適化

問題:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{：目的関数}$$

$$\text{s.t.} \quad g_k(\boldsymbol{x}) \leq 0 \quad (k = 1,2,\cdots,m) \qquad (2) \quad \text{：制約}$$

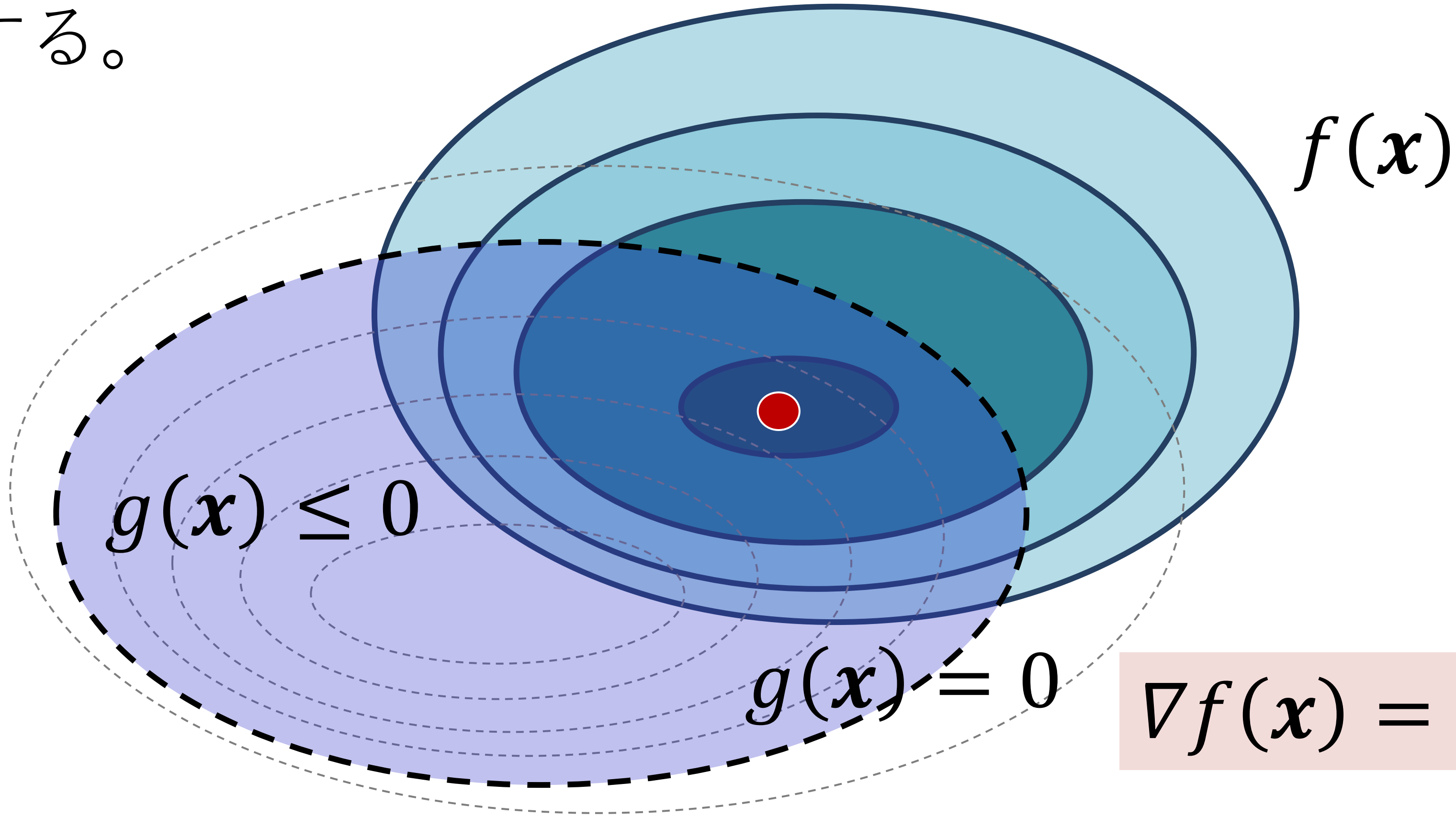
ここで,

$\boldsymbol{x} \in R^n$ ：設計変数

$f(\boldsymbol{x}), g(\boldsymbol{x})$ は, $R^n \rightarrow R$の関数

Case I:　非アクティブな制約

制約がないときの最適解●が実行可能解(feasible region)に属する。

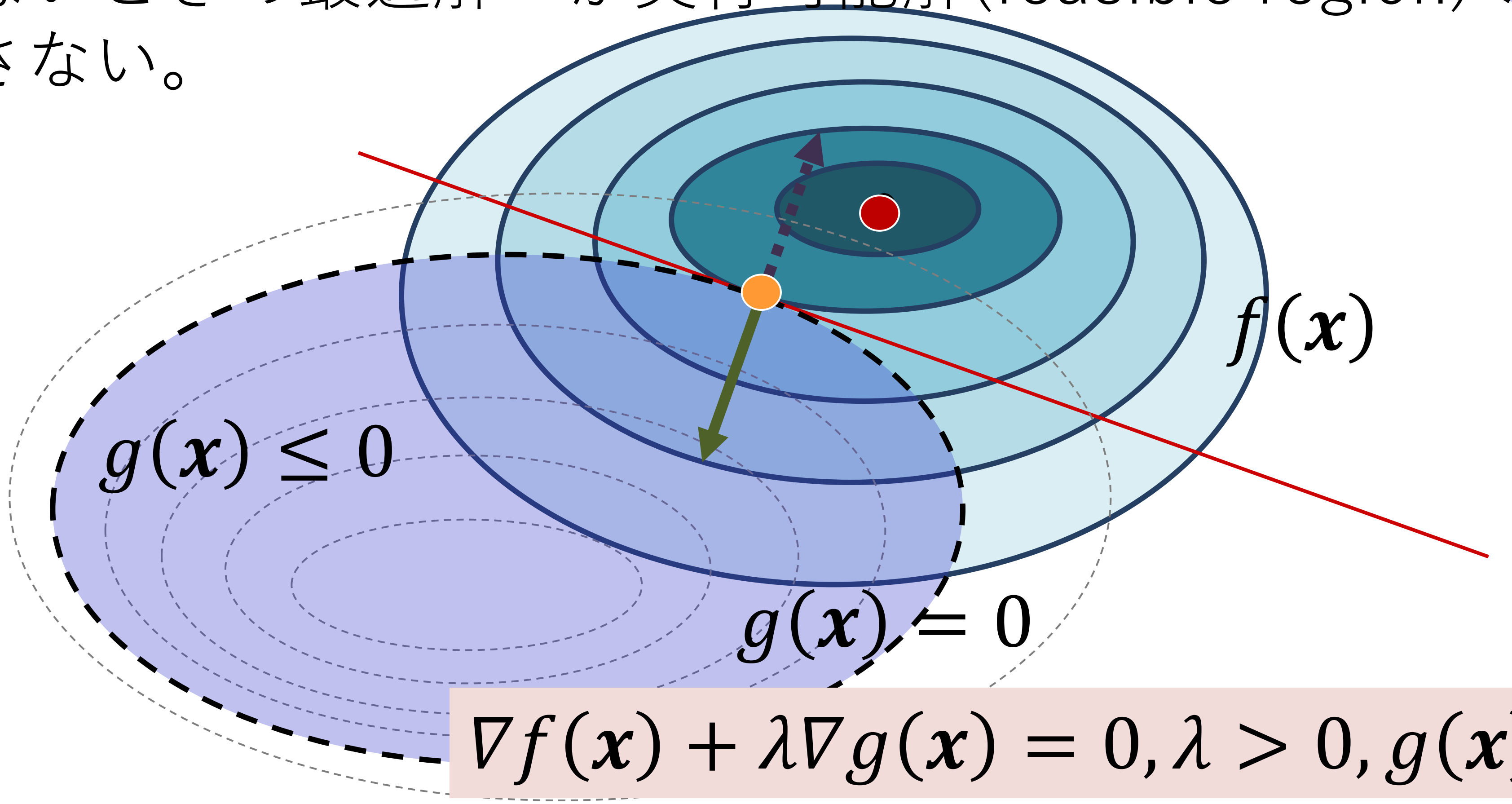


$$\nabla f(\boldsymbol{x}) = 0, \quad g(\boldsymbol{x}) \leq 0$$

(問題は，制約がないときの最適化問題と同じ)

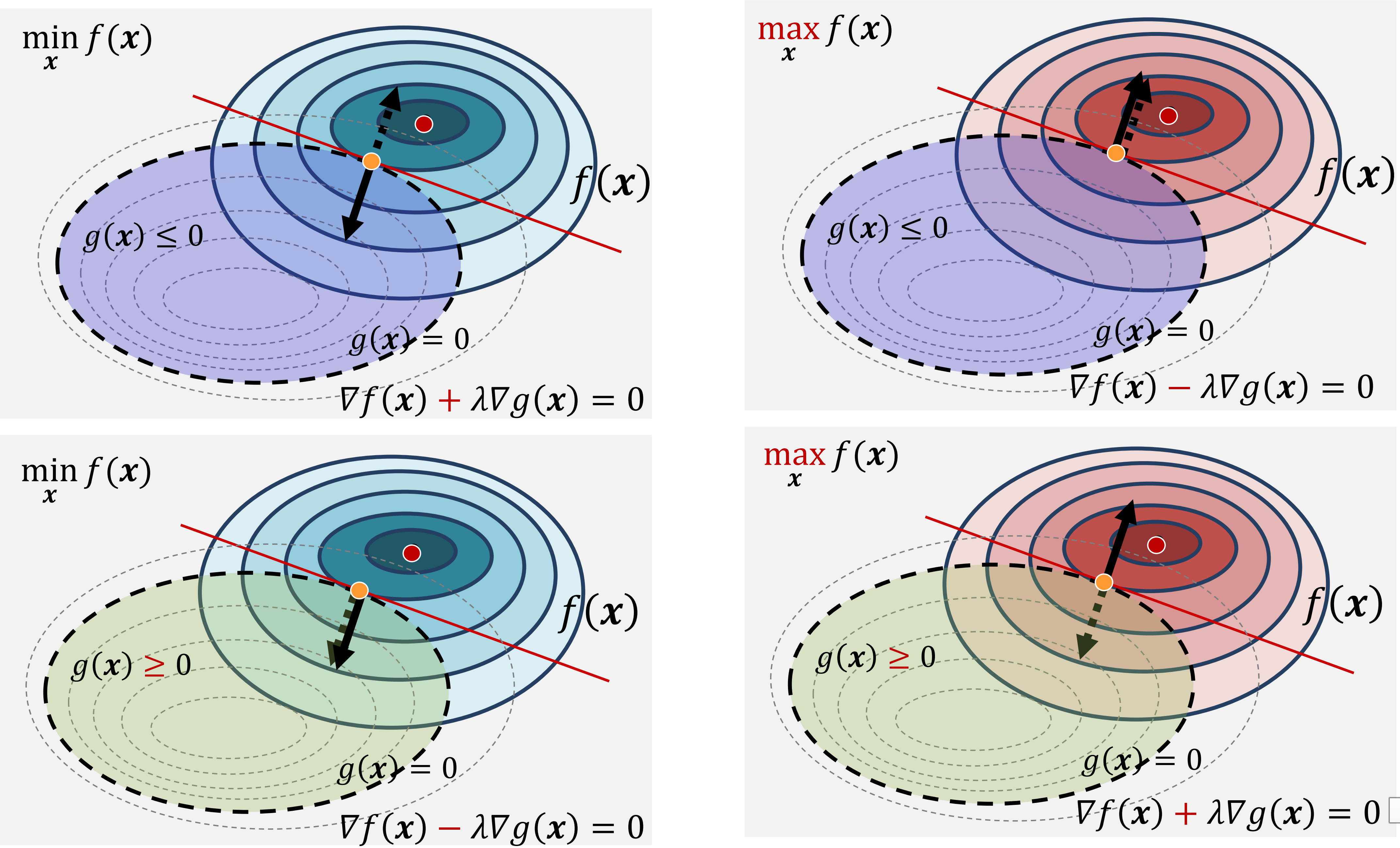
Case II:アクティブな制約

制約がないときの最適解●が実行可能解(feasible region)⬭
に属さない。

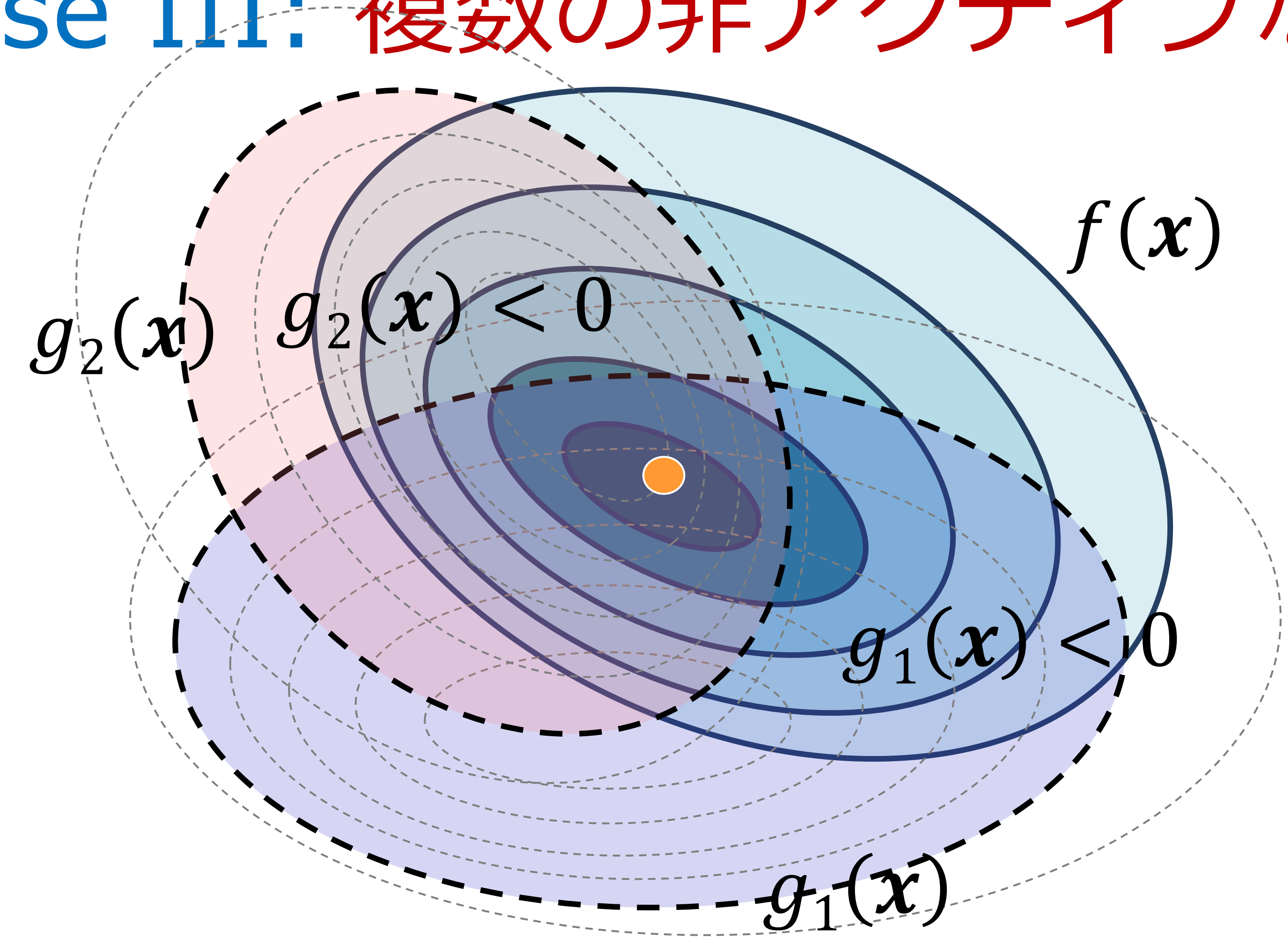


$$\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0, \lambda > 0, g(\boldsymbol{x}) = 0$$

(問題は，等式制約の最適化問題と同じ)

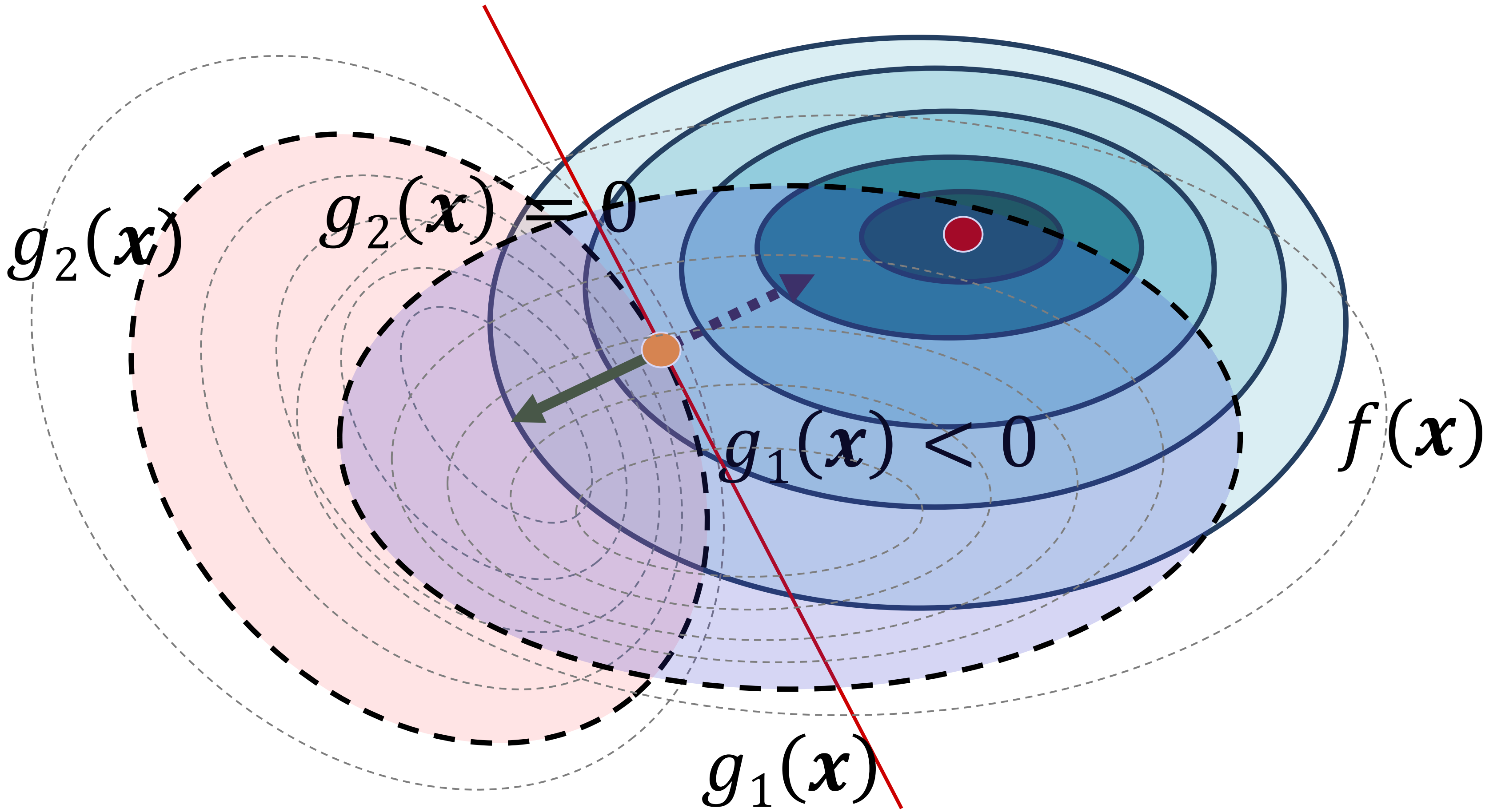

$\min_{\boldsymbol{x}} f(\boldsymbol{x})$
$f(\boldsymbol{x})$
$g(\boldsymbol{x}) \leq 0$
$g(\boldsymbol{x}) = 0$
$\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0$
$\max_{\boldsymbol{x}} f(\boldsymbol{x})$
$f(\boldsymbol{x})$
$g(\boldsymbol{x}) \leq 0$
$g(\boldsymbol{x}) = 0$
$\nabla f(\boldsymbol{x}) - \lambda \nabla g(\boldsymbol{x}) = 0$
$\min_{\boldsymbol{x}} f(\boldsymbol{x})$
$f(\boldsymbol{x})$
$g(\boldsymbol{x}) \geq 0$
$g(\boldsymbol{x}) = 0$
$\nabla f(\boldsymbol{x}) - \lambda \nabla g(\boldsymbol{x}) = 0$
$\max_{\boldsymbol{x}} f(\boldsymbol{x})$
$f(\boldsymbol{x})$
$g(\boldsymbol{x}) \geq 0$
$g(\boldsymbol{x}) = 0$
$\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0$

Case III: 複数の非アクティブな制約



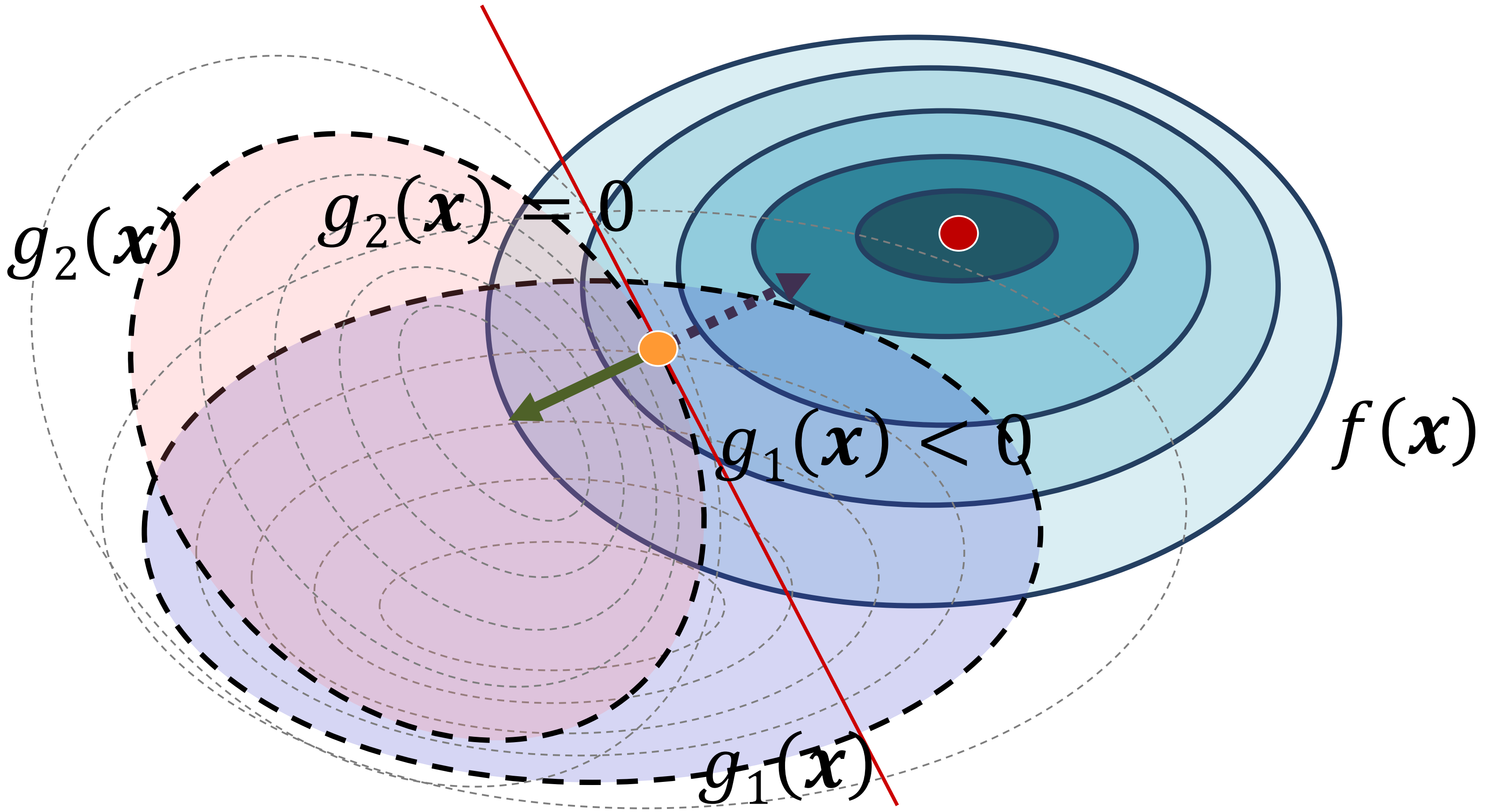
$$\nabla f(\boldsymbol{x}) = 0, g_1(\boldsymbol{x}) < 0, g_2(\boldsymbol{x}) < 0$$

Case IV:非アクティブとアクティブな制約の混合



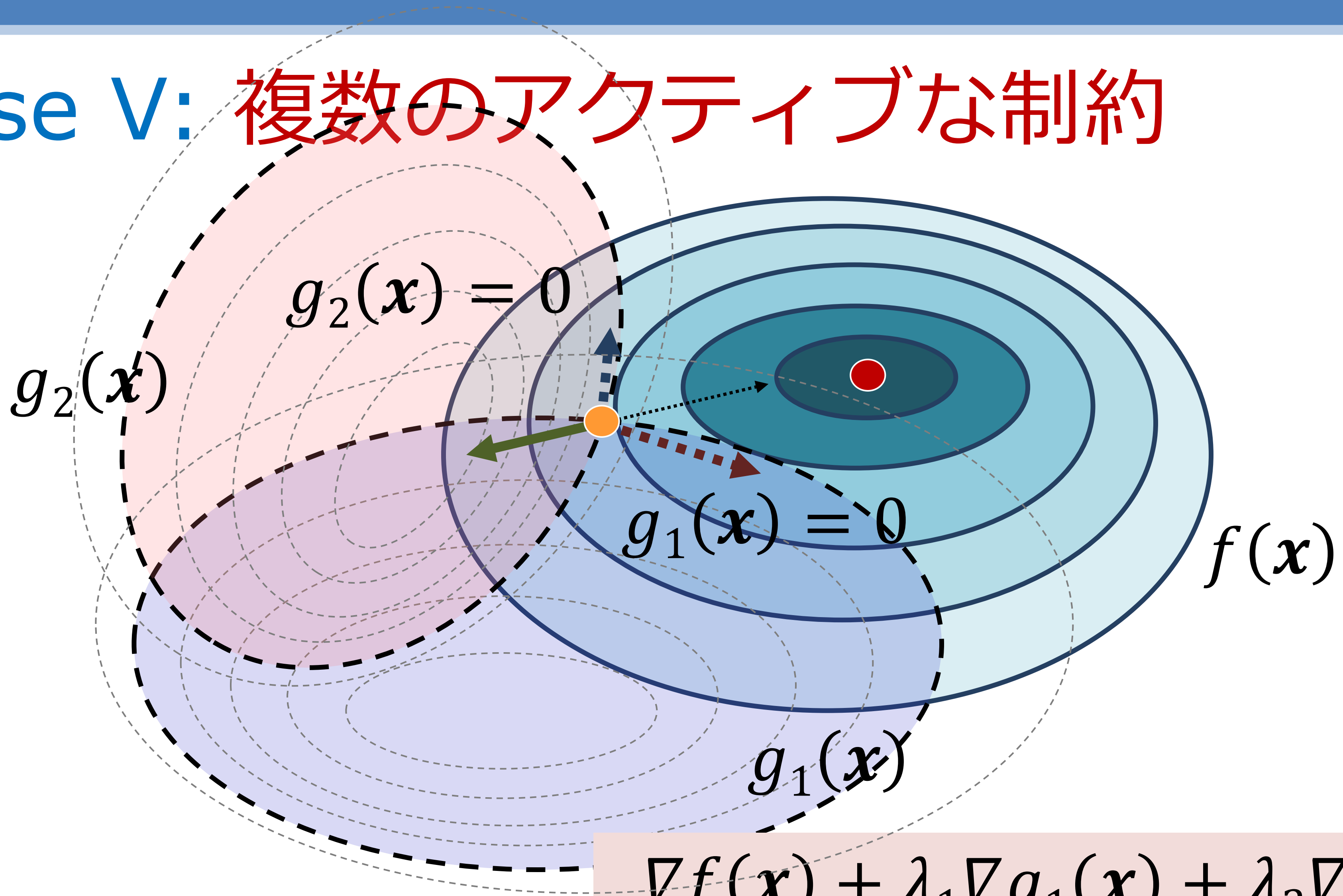
$$\nabla f(\boldsymbol{x}) + \lambda \nabla g_2(\boldsymbol{x}), = 0 \quad g_1(\boldsymbol{x}) < 0,\ g_2(\boldsymbol{x}) = 0$$

Case IV:非アクティブとアクティブな制約の混合



$$\nabla f(\boldsymbol{x}) + \lambda \nabla g_2(\boldsymbol{x}), = 0 \quad g_1(\boldsymbol{x}) < 0,\ g_2(\boldsymbol{x}) = 0$$

Case V: 複数のアクティブな制約



$$\nabla f(\boldsymbol{x}) + \lambda_1 \nabla g_1(\boldsymbol{x}) + \lambda_2 \nabla g_2(\boldsymbol{x}) = \mathbf{0}$$
$$g_1(\boldsymbol{x}) = g_2(\boldsymbol{x}) = 0$$

アクティブ／非アクティブな制約

Case I:　制約があってもなくても解に影響を与えない
⇒　非アクティブな制約

$$\nabla f(\boldsymbol{x}) = 0, \quad g(\boldsymbol{x}) < 0. \tag{3}$$

Case II: 制約があることで，解が変わる
⇒　アクティブな制約

$$\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0, g(\boldsymbol{x}) = 0. \tag{4}$$

非アクティブな制約に対するラグランジュ定数λ_kは，$\lambda_k = 0$ とする。このとき，解は次式を満たす。

$$\nabla f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k \nabla g_k(\boldsymbol{x}) = 0, \qquad (5)$$
$$\boxed{\lambda_k}\boxed{g_k(\boldsymbol{x})} = 0, \lambda_k \geq 0$$

(Kuhn-Tucker条件)

制約が非アクティブのとき:

$$\lambda_k = 0$$

制約がアクティブのとき:

$$g_k(\boldsymbol{x}) = 0$$

Lagrangeの未定乗数法

Lagrange関数

$$L(x,\boldsymbol{\lambda}) = f(x) + \sum_{k=1}^{m} \lambda_k g_k(x)$$

とおくとき，最適解はKuhn-Tucker の方程式を満たす。

$$\nabla L(x,\boldsymbol{\lambda}) = \nabla f(x^*) + \sum_{k=1}^{m} \lambda_k^* \nabla g_k(x^*) = 0$$

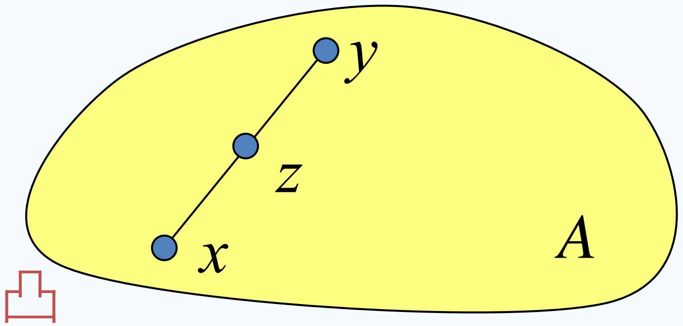
$$\lambda_k^* g_k(x^*) = 0, \qquad k = 1,2,\cdots,m.$$

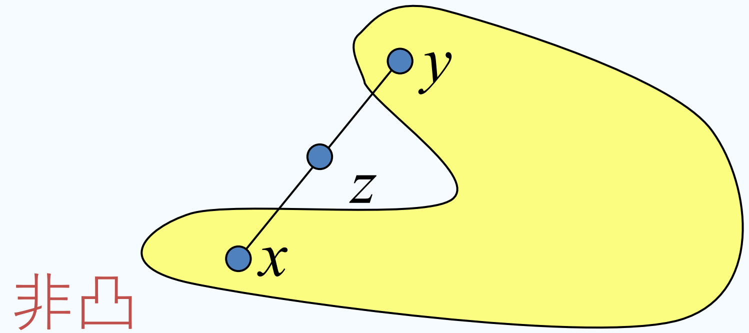
$$\lambda_k^* \geq 0, \qquad k = 1,2,\cdots,m.$$

凸計画問題に対して，KT 方程式は最適解を得るための必要十分条件となる。

凸集合, 凸関数, 凸計画問題

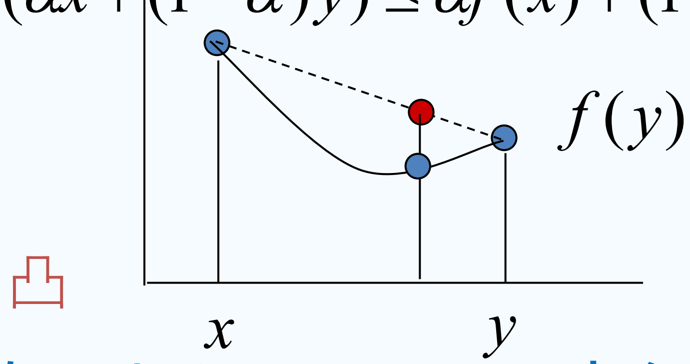
集合Aの任意の2点 $\boldsymbol{x}, \boldsymbol{y}$ の線分上の点$\boldsymbol{z}$が集合Aに含まれるとき,
集合Aを凸集合と呼ぶ。

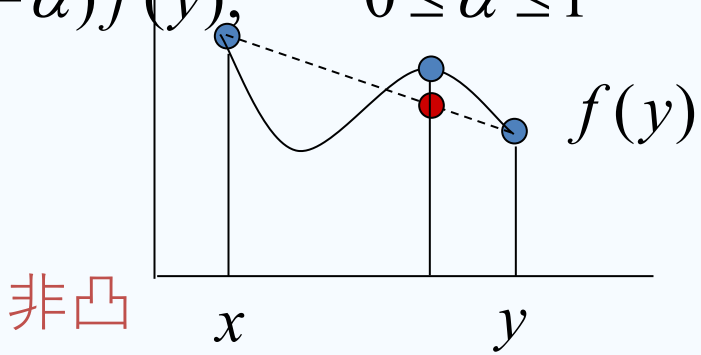




任意の$\boldsymbol{x}, \boldsymbol{y}$について次式が成立するとき， f を凸関数と呼ぶ。

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y), \qquad 0 \leq \alpha \leq 1$$





$f(x)$が凸関数であり，さらに実行可能領域が凸集合であるような数理計画問題のことを凸計画問題と呼ぶ。

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \qquad \text{s.t.} \quad g_k(\boldsymbol{x}) \le 0, \quad k = 1,2,\cdots,m. \tag{2}$$

元々の最適化問題

$\Leftrightarrow$

$$\nabla f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k \nabla g_k(\boldsymbol{x}) = 0, \qquad \lambda_k g_k(\boldsymbol{x}) = 0, \lambda_k > 0 \tag{5}$$

Kuhn-Tucker 条件

$\Leftrightarrow$

$$\min_{\boldsymbol{x},\boldsymbol{\lambda}} L(\boldsymbol{x},\boldsymbol{\lambda}); \quad L(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}) \tag{6}$$

$$\lambda_k g_k(\boldsymbol{x}) = 0, \lambda_k > 0$$

ラグランジュの未定乗数法

§ 双対問題

ラグランジュ関数は，設計変数 $\boldsymbol{x}$ とラグランジュ定数 $\boldsymbol{\lambda}$ の関数

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x})$$

であり，また $g_k(\boldsymbol{x})$ は複雑な関数であることが多い。
このため，一般に最適化は難しい。

$L(\boldsymbol{x}, \boldsymbol{\lambda})$の最小化（最大化）問題と等価な $\boldsymbol{\lambda}$ だけの関数 $\phi(\boldsymbol{\lambda})$ の最大化（最小化）問題＝双対問題＝が存在し，より簡単に解くことができる。

双対問題

問題:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \qquad \text{s.t.} \quad g_k(\boldsymbol{x}) \le 0, \quad k = 1,2,\cdots,m.$$

ラグランジュ関数 (Lagrangian):

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}), \qquad \lambda_k \ge 0$$

$\boldsymbol{x}^*, \boldsymbol{\lambda}^*$ を $\boldsymbol{x}, \boldsymbol{\lambda}$ の最適解とすると,

$$L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = f(\boldsymbol{x}^*) \le f(\boldsymbol{x}), \quad \{\boldsymbol{x} | g_k(\boldsymbol{x}) \le 0, k = 1,2,\cdots,m\} \quad (7)$$

双対問題

ラグランジュ関数:

$$L(\boldsymbol{x}^*, \boldsymbol{\lambda}) = f(\boldsymbol{x}^*) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}^*)$$

定義により　$\lambda_k \geq 0, g_k(\boldsymbol{x}^*) \leq 0$ であるから,

$$\lambda_k g_k(\boldsymbol{x}^*) \leq 0$$

よって,

$$L(\boldsymbol{x}^*, \boldsymbol{\lambda}) \leq f(\boldsymbol{x}^*) \tag{8}$$

双対問題

$\phi(\boldsymbol{\lambda})$ を $L(\boldsymbol{x},\boldsymbol{\lambda})$ の$\boldsymbol{x}$ での最小値とする。

$$\phi(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}} L(\boldsymbol{x},\boldsymbol{\lambda})$$

このとき,

$$\phi(\boldsymbol{\lambda}) \leq L(\boldsymbol{x}^*,\boldsymbol{\lambda}) \tag{9}$$

よって,

$$\phi(\boldsymbol{\lambda}) \leq f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$$

$$\underbrace{\max_{\boldsymbol{\lambda}} \phi(\boldsymbol{\lambda}) = f(\boldsymbol{x}^*)}_{\text{双対問題}} \overbrace{= \min_{\boldsymbol{x}} f(\boldsymbol{x})}^{\text{主問題}} \tag{10}$$

双対問題による最小化の問題の解き方

1. ラグランジュ定数$\boldsymbol{\lambda}$を導入してラグランジュ関数を定義する。

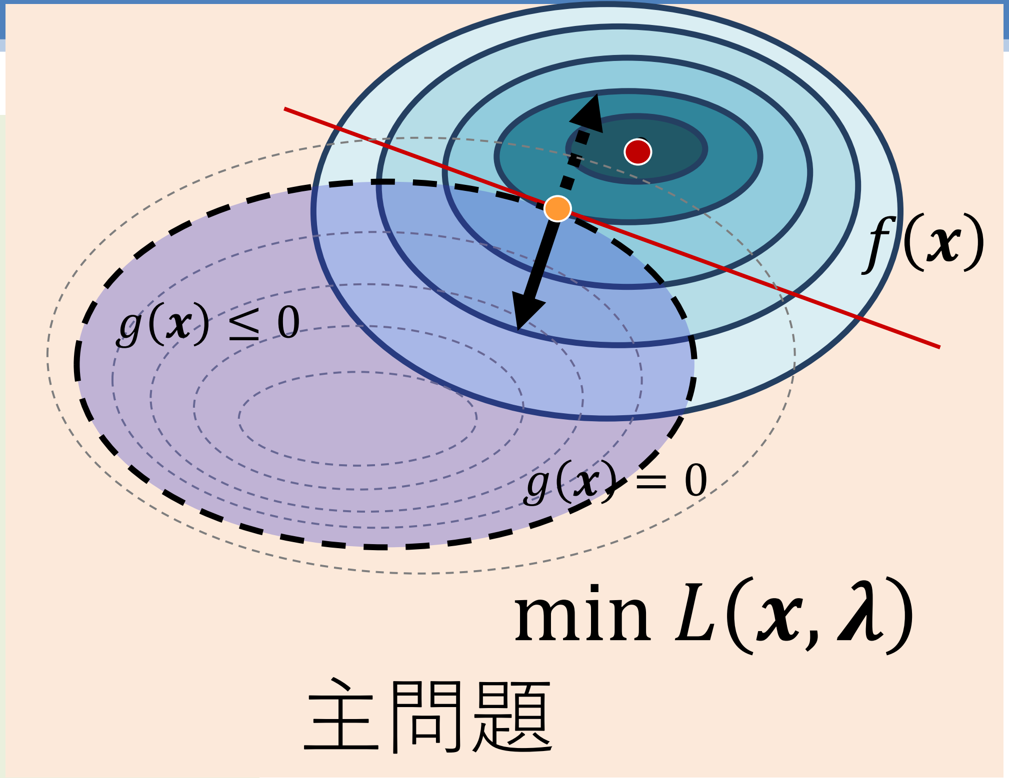
$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}),$$

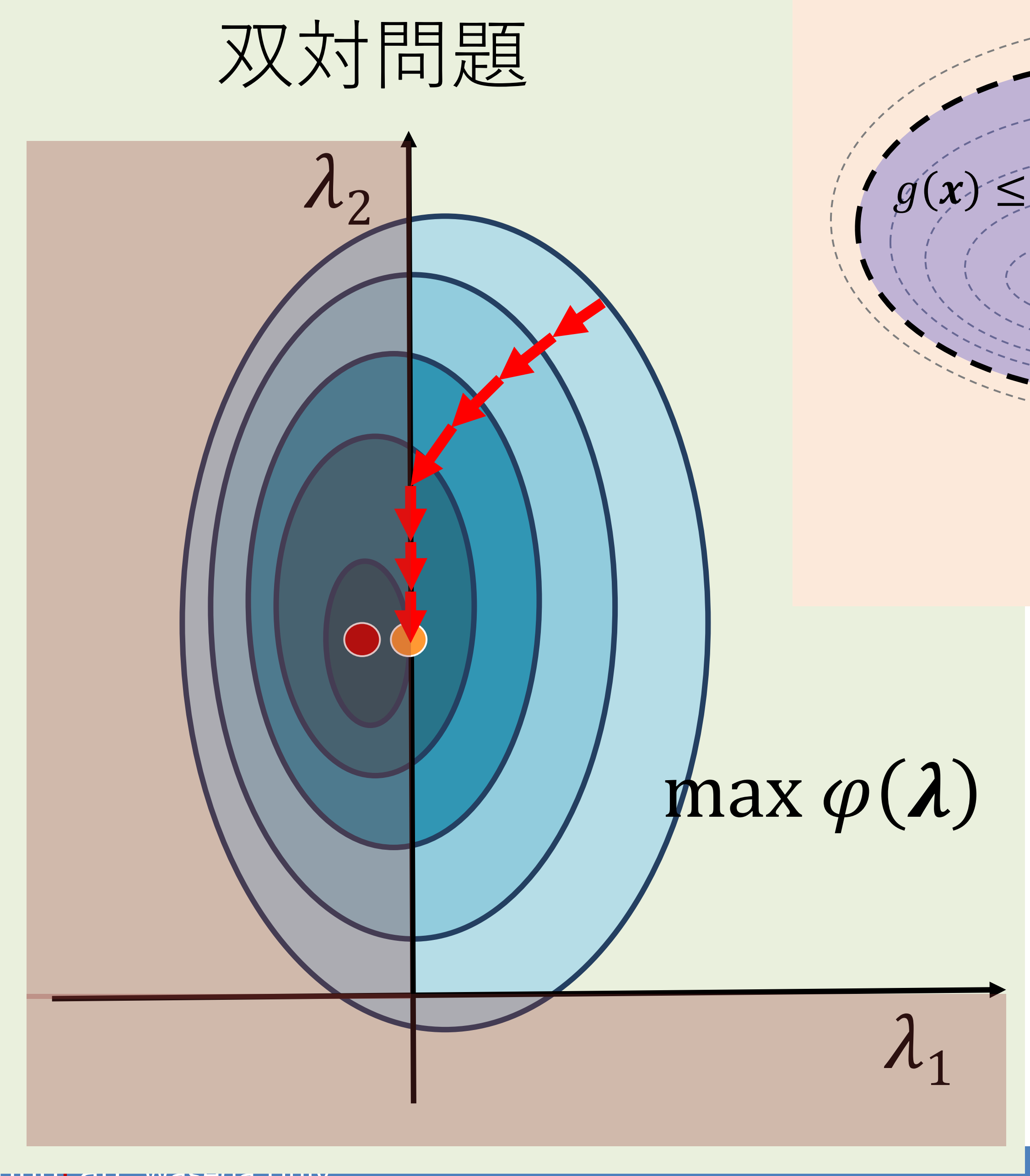
2. $L(\boldsymbol{x}, \boldsymbol{\lambda})$ を $\boldsymbol{x}$ で最小化して $\phi(\boldsymbol{\lambda})$ を求める。($\frac{\partial}{\partial x} L(\boldsymbol{x}, \lambda) = 0$ の結果を $L(\boldsymbol{x}, \lambda)$に代入して$\boldsymbol{x}$ を消去する).

$$\phi(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda})$$

3. $\boldsymbol{\lambda}$, $\lambda_k \geq 0, k = 1,2, \cdots, m$ で，　$\phi(\boldsymbol{\lambda})$を最大化する。

$$f(\boldsymbol{x}^*) = \max_{\boldsymbol{\lambda}} \phi(\boldsymbol{\lambda}), \lambda_k \geq 0, k = 1,2, \cdots, m$$





多くの場合
双対問題が
より簡単

例題

条件

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 4\begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{x} + 7 \leq 0$$

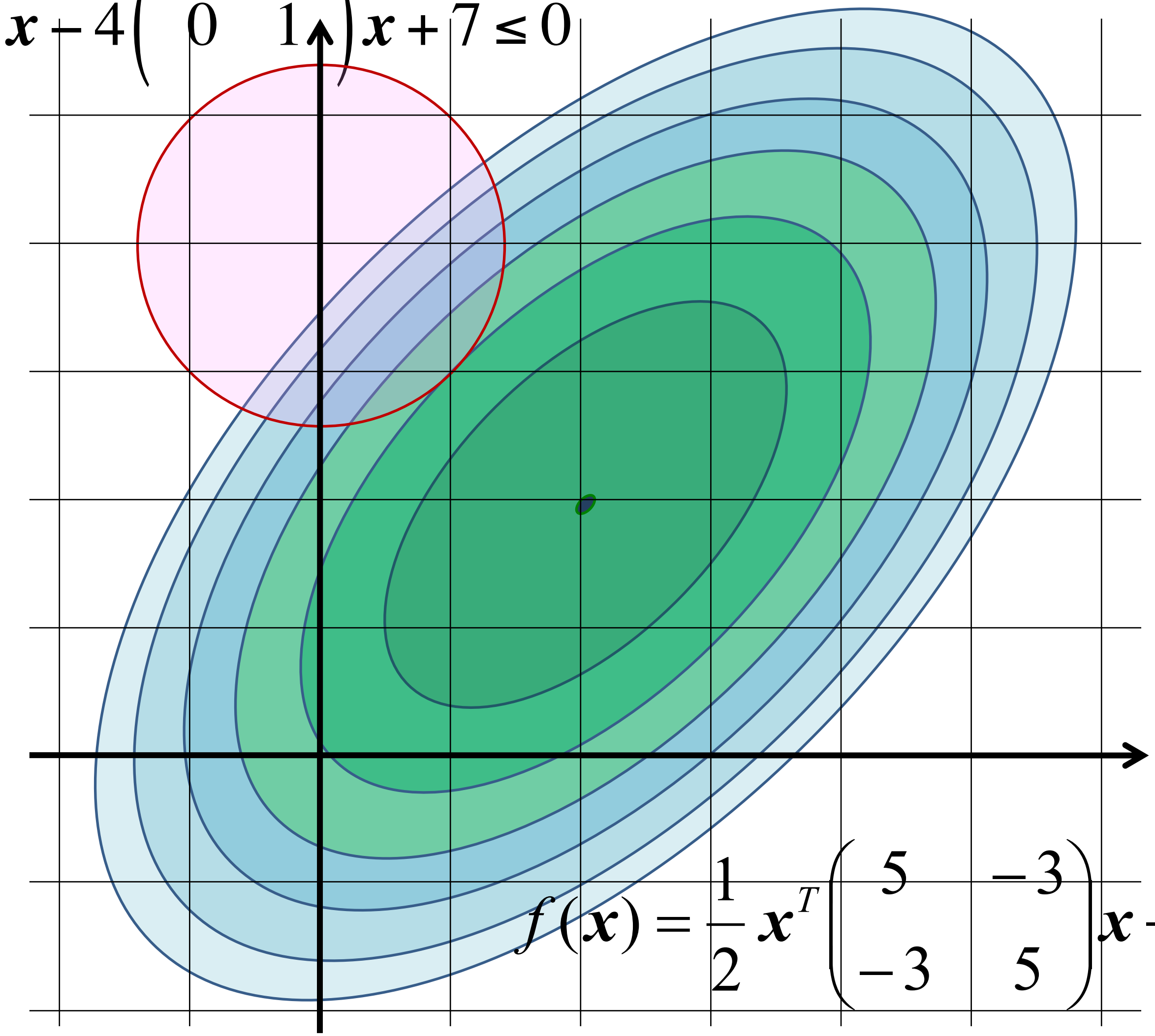
の下で，

関数，

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T\boldsymbol{x}$$

の最小値を与える$\boldsymbol{x}$を求めよ。

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 4\begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{x} + 7 \leq 0$$



$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T\boldsymbol{x} = c$$

評価関数を次のように置く

$$
\begin{aligned}
L(\boldsymbol{x},\lambda) &= f(\boldsymbol{x}) + \lambda g(\boldsymbol{x}) \\
&= \frac{1}{2}\boldsymbol{x}^T \begin{pmatrix} 5 & -3 \\ -3 & 4 \end{pmatrix} \boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \boldsymbol{x} + \lambda\left(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 4\begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{x} + 7 \right) \\
&= \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} + \lambda\left(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \boldsymbol{c}^T\boldsymbol{x} + d \right) \\
&= \frac{1}{2}\boldsymbol{x}^T (A + \lambda I)\boldsymbol{x} - (\boldsymbol{b} + \boldsymbol{c})^T \boldsymbol{x} + \lambda d
\end{aligned}
$$

$$
A = \begin{pmatrix} 5 & -3 \\ -3 & 4 \end{pmatrix}, \boldsymbol{b} = 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{c} = 4\begin{pmatrix} 0 \\ 1 \end{pmatrix}, d = 7
$$

Lを$\boldsymbol{x}$で偏微分して0と置く。

$$\frac{\partial}{\partial \boldsymbol{x}} L(\boldsymbol{x}, \lambda) = (A + \lambda I)\boldsymbol{x} - (b + \lambda c) = 0$$

$$\boldsymbol{x} = (A + \lambda I)^{-1}(b + \lambda c)$$

A,b,c,dに具体的数値を戻し， $4\eta = \lambda$ とおいて整理すると，

$$\boldsymbol{x} = \begin{pmatrix} 5+\lambda & -3 \\ -3 & 5+\lambda \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 4+4\lambda \end{pmatrix}$$

$$= \frac{4}{(\lambda+2)(\lambda+8)} \begin{pmatrix} 5+\lambda & 3 \\ 3 & 5+\lambda \end{pmatrix} \begin{pmatrix} 1 \\ 1+\lambda \end{pmatrix}$$

$$= \frac{4}{(\lambda+8)} \begin{pmatrix} 4 \\ \lambda+4 \end{pmatrix} = \frac{4}{(\eta+2)} \begin{pmatrix} 1 \\ \eta+1 \end{pmatrix}$$

L　に代入し，$\boldsymbol{x}$を消去して，双対問題を導く。

$$\frac{L(\eta)}{4} = -\frac{2}{2+\eta}\begin{pmatrix} 1 \\ 1+4\eta \end{pmatrix}^T \begin{pmatrix} 1 \\ 1+\eta \end{pmatrix} + 7\eta$$

$$= \frac{-1}{2+\eta}\left(\eta^2 - 4\eta + 4\right) = \frac{-1}{2+\eta}(\eta-2)^2$$

これを　η　で最大化する。

$$\frac{\partial}{\partial\eta}\frac{L(\eta)}{4} = \frac{\partial}{\partial\eta}\frac{-1}{\eta+2}(\eta-2)^2 = -\frac{(\eta-2)(\eta+6)}{(\eta+2)^2}$$

$\eta>0$　で，上式を最大化する　η　と，$f(x)$の最小値は，

$$\eta = 2$$

$$\min_{\boldsymbol{x}} f(x) = \min_{\boldsymbol{x},\lambda} L(x,\lambda) = \max_{\eta} L(\eta) = L(2) = 0$$

まとめ

- $\min_{\boldsymbol{x}} f(\boldsymbol{x})$, s.t. $g(\boldsymbol{x}) \leq 0$ の問題は，
  $L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}), \lambda_k \geq 0$ の最小化（ラグランジュの未定乗数法）
  $\nabla f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k \nabla g_k(\boldsymbol{x}) = 0, \lambda_k g_k(\boldsymbol{x}) = 0, \lambda_k > 0$ (KT条件）
  により解く。
- $L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x}), \lambda_k \geq 0$ の最小化は，双対問題の最大化によって解く。

演習問題

1.

条件

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \begin{pmatrix} 2 & 2 \end{pmatrix}\boldsymbol{x} + 2 \leq 0 \qquad (1)$$

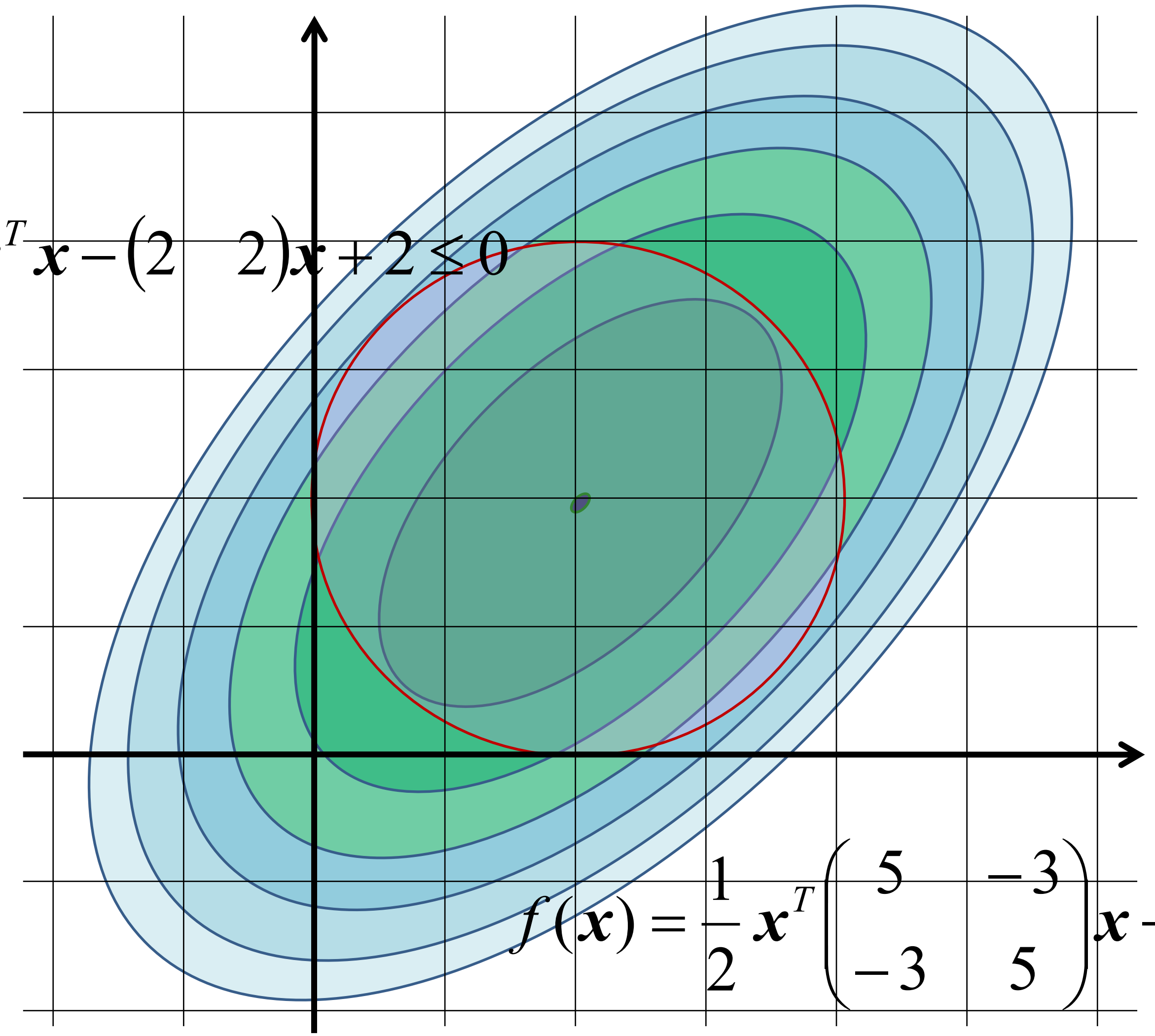
の下で，

関数，

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T\boldsymbol{x} \qquad (2)$$

の最小値を与えるxを求めよ。

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \begin{pmatrix}2 & 2\end{pmatrix}\boldsymbol{x} + 2 \le 0$$



$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix}5 & -3\\ -3 & 5\end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix}1\\ 1\end{pmatrix}^T\boldsymbol{x} = c$$

演習問題

2. SVMの説明にある，式(4)(5)(6)を実際に計算して導け。

3. サポートベクタに対するラグランジュ係数を　α^0マージンを
　ρ　とするとき，

$$(\boldsymbol{\psi}^0 * \boldsymbol{\psi}^0) = \sum_{i=1}^{m} \alpha_i^0$$

$$\rho(\boldsymbol{\psi}^0) = \frac{1}{|\boldsymbol{\psi}^0|}$$

$$L(\boldsymbol{\alpha}) > L(\boldsymbol{\alpha}^0) = \frac{1}{2}(\boldsymbol{\psi}^0 * \boldsymbol{\psi}^0) = \frac{1}{2}\frac{1}{\rho(\boldsymbol{\psi}^0)^2}, \qquad \boldsymbol{\alpha} \neq \boldsymbol{\alpha}^0$$

を証明せよ。

演習問題

4. カテゴリAとして，(2,2),(0,1),(-1,2)が， カテゴリBとして，(1,-1),(-1,-2),(-3,0) が与えられたとき，これらのカテゴリをマージン最大化基準で分離する直線を求めよ。