

関数の勾配とヘッセ行列

$\nabla f(\mathbf{x})$: 点 \mathbf{x} における関数 f の勾配

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

1. (1)

$$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$$

$$\frac{\partial f(\mathbf{x})}{\partial x_1} = 6x_1 - 2x_2 + 6$$

$$\frac{\partial f(\mathbf{x})}{\partial x_2} = -2x_1 + 6x_2 - 10$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix}$$

1. (2) 点 $\mathbf{a}=(0,0)^T$ における関数 f の勾配

$$\nabla f(\mathbf{a}) = \begin{bmatrix} 6 \\ -10 \end{bmatrix}$$

$\nabla^2 f(\mathbf{x})$: 点 \mathbf{x} における関数 f のヘッセ行列

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}$$

対称行列

1. (3)

$$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix}$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$$

固有多項式

$$g_A(t) = |tE - A| \quad A: \text{正方行列}$$

行列Aの固有値: $g_A(t) = 0$ の根 (複素根も含む)

$$A = \nabla^2 f(x) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$$

$$|tE - A| = \begin{vmatrix} t-6 & 2 \\ 2 & t-6 \end{vmatrix} = t^2 - 12t + 32 = (t-4)(t-8)$$

$\lambda = 4, 8$

$$T_A(x) = Ax \rightarrow Ax = \lambda x \rightarrow Ax = \lambda Ex$$

$$(\lambda E - A)x = 0 \quad x: \text{固有ベクトル}$$

$$\lambda = 4 \text{ のとき } \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} x = 0 \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda = 8 \text{ のとき } \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} x = 0 \quad x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

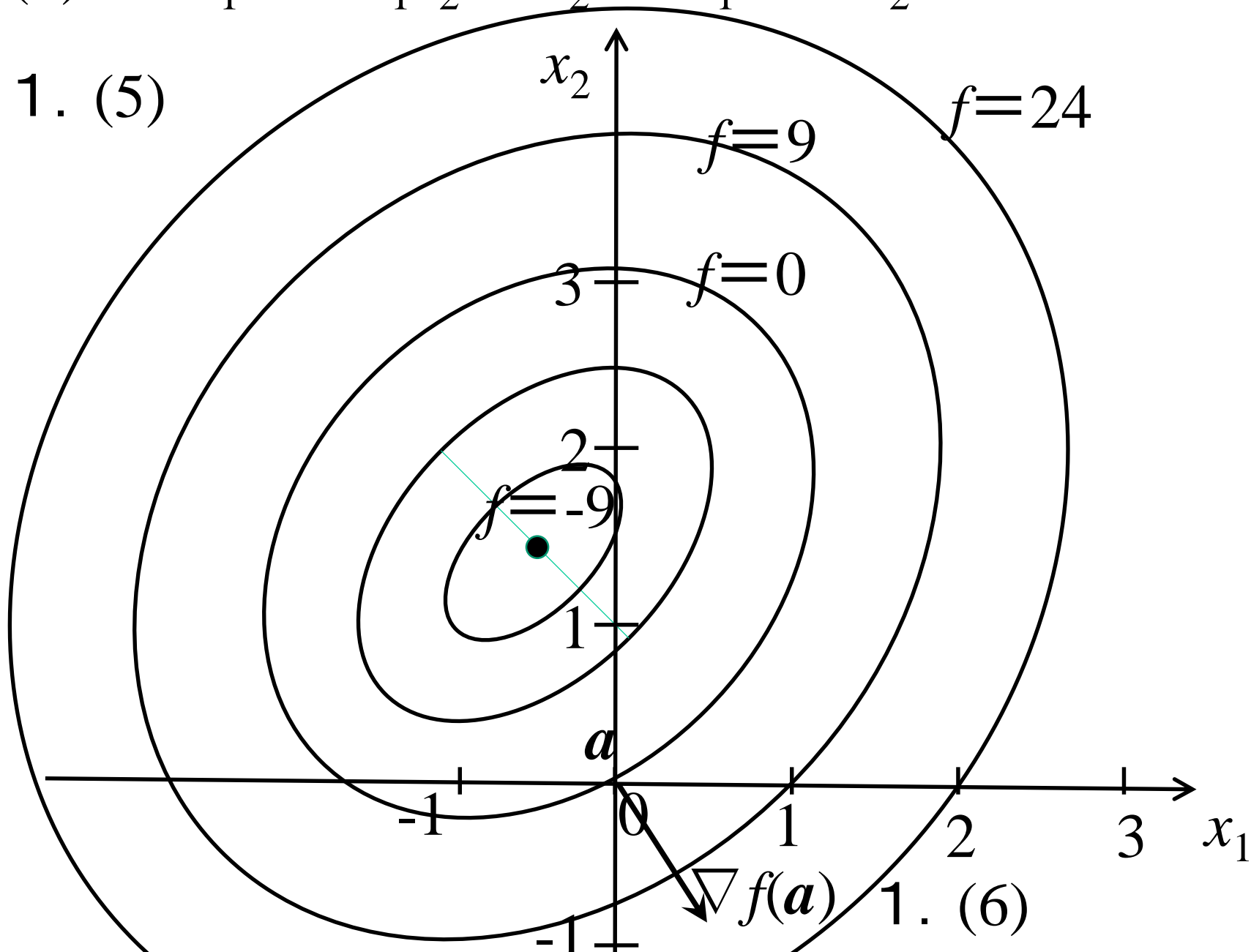
1. (4)

固有値： $\lambda_1=4, \lambda_2=8$

固有ベクトル： $x_1=(1,1)^T, x_2=(1,-1)^T$

$f(x) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$ の等高線

1. (5)



1. (6)

制約なし問題の最適性条件

$$\nabla f(x^*) = \mathbf{0}$$

x^* が局所的最適解であるための必要条件

x^* : 関数 f の停留点

1次の必要条件

1. (7)

最適性の1次の必要条件

$$\nabla f(\mathbf{x}) = \mathbf{0}$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix} = \mathbf{0}$$

$$\mathbf{x}^* = (-1/2, 3/2)$$

A: 半正定値行列

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0 \quad (\text{すべての } \mathbf{x} \text{ について})$$

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

$\nabla^2 f(\mathbf{x}^*)$ は半正定値

最適性の2次の
必要条件

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

$\nabla^2 f(\mathbf{x}^*)$ は正定値

最適性の2次の
十分条件

1. (8)

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$$

固有値: $\lambda_1=4, \lambda_2=8$

固有値がすべて正である対称行列



正定値行列

$\nabla^2 f(\mathbf{x}^*)$ は正定値

\mathbf{x}^* は最適性の2次の必要条件と
十分条件を満たす

最急降下法

<最急降下法>

(0) 出発点 $x^{(0)}$ を選び, $k:=0$ とおく.

(1) $\nabla f(x^{(k)})=0$ ならば計算終了. さもなければ
 $d^{(k)} := -\nabla f(x^{(k)})$ とおいてステップ(2)へ.

(2) ステップ幅 $\alpha^{(k)}$ を求め, 次の点
 $x^{(k+1)} := x^{(k)} + \alpha^{(k)} d^{(k)}$ を定める.

$k:=k+1$ とおいてステップ(1)へ戻る.

2. (1)

$$\mathbf{d} = -\nabla f(\mathbf{x})$$

$$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix}$$

$$\mathbf{x}^{(0)} = (0, 0)^T \quad \text{のとき}$$

$$\mathbf{d}^{(0)} = (-6, 10)^T$$

$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$ の等高線

2. (2) $f=24$

$f=9$

$\nabla f(\mathbf{x}) =$

$$\begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix}$$

$f=0$

3

2

$f=-9$

$\mathbf{x}^{(1)}$

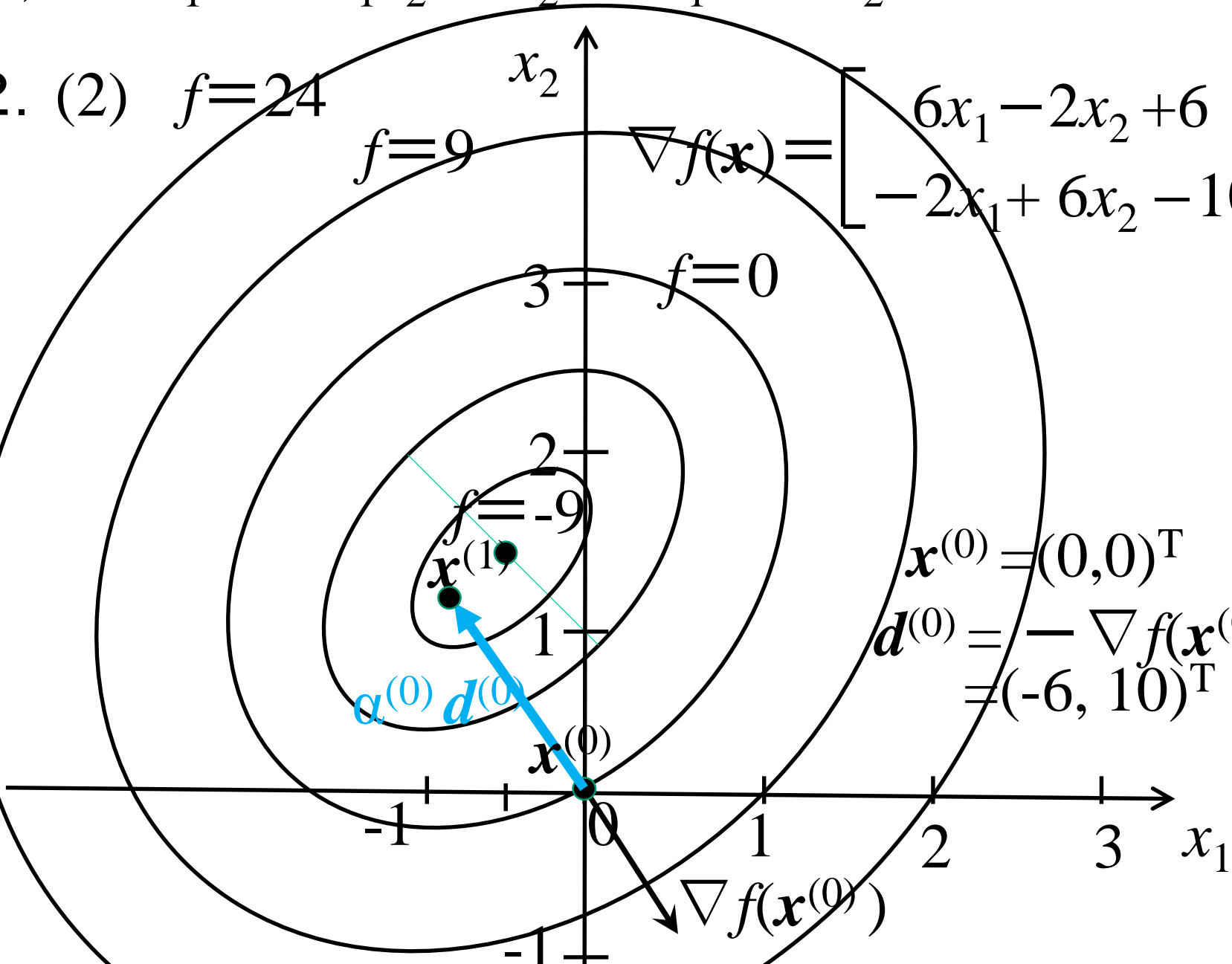
$\alpha^{(0)}$ $d^{(0)}$

$\mathbf{x}^{(0)}$

$\mathbf{x}^{(0)} = (0, 0)^T$

$d^{(0)} = -\nabla f(\mathbf{x}^{(0)})$
 $= (-6, 10)^T$

$\nabla f(\mathbf{x}^{(0)})$



ニュートン法と準ニュートン法

<ニュートン法>

(0) 出発点 $x^{(0)}$ を選び, $k:=0$ とおく.

(1) $\nabla f(x^{(k)})=0$ ならば計算終了. さもなければ

$$\nabla^2 f(x^{(k)})d = -\nabla f(x^{(k)})$$

の解 $d^{(k)}$ を求め, ステップ(2)へ.

(2) 次の点を $x^{(k+1)} := x^{(k)} + d^{(k)}$ とする.

$k:=k+1$ とにおいてステップ(1)へ戻る.

$$2. (3) \quad \nabla^2 f(\mathbf{x}) \mathbf{d} = -\nabla f(\mathbf{x})$$

$$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 6 \\ -2x_1 + 6x_2 - 10 \end{bmatrix}$$

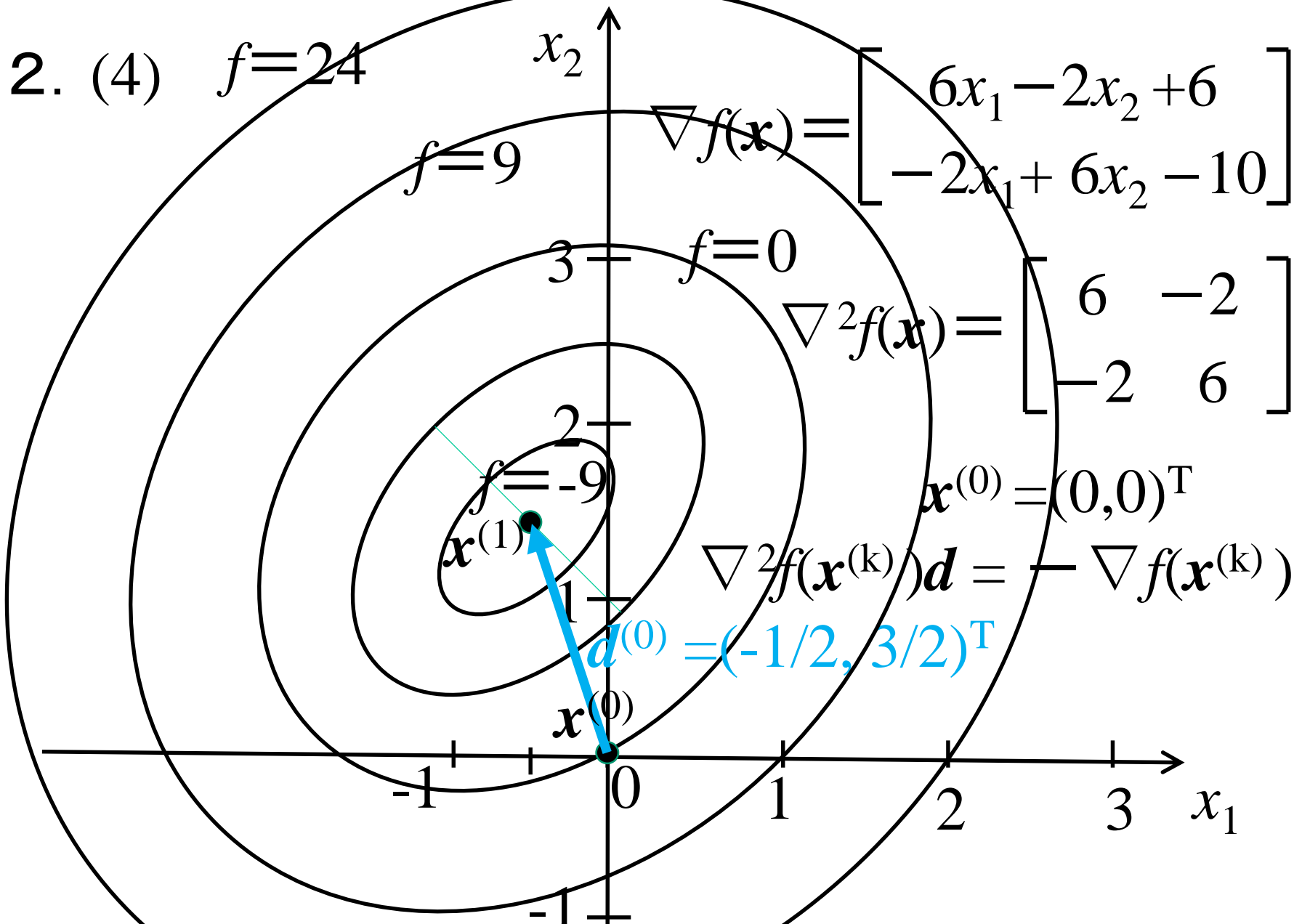
$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$$

$\mathbf{x} = (0, 0)^T$ のとき

$$\begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \mathbf{d} = \begin{bmatrix} -6 \\ 10 \end{bmatrix} \longrightarrow \mathbf{d} = (-1/2, 3/2)^T$$

$f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + 6x_1 - 10x_2$ の等高線

2. (4) $f=24$



$$f(\mathbf{x}) = (x_1 - 1)^2 + 10(x_1^2 - x_2)^2$$

3. (1)

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 1) + 40(x_1^2 - x_2)x_1 \\ -20(x_1^2 - x_2) \end{bmatrix}$$

3. (2)

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 2 + 120x_1^2 - 40x_2 & -40x_1 \\ -40x_1 & 20 \end{bmatrix}$$

3. (3)

$$\mathbf{x}^* = (1, 1)^T \quad \text{のとき} \quad \nabla^2 f(\mathbf{x}^*) = \begin{bmatrix} 82 & -40 \\ -40 & 20 \end{bmatrix}$$

3. (4)

$$A = \nabla^2 f(\mathbf{x}) = \begin{bmatrix} 82 & -40 \\ -40 & 20 \end{bmatrix}$$

$$|tE - A| = \begin{vmatrix} t-82 & 40 \\ 40 & t-20 \end{vmatrix} = t^2 - 102t + 40 = 0$$

$\lambda = 101.6, 0.4$

条件数 $\gamma = \lambda_{\max} / \lambda_{\min} = 101.6 / 0.4 \doteq 254$

3. (5)

$$\mathbf{d} = -\nabla f(\mathbf{x})$$

$$f(\mathbf{x}) = (x_1 - 1)^2 + 10(x_1^2 - x_2)^2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 1) + 40(x_1^2 - x_2)x_1 \\ -20(x_1^2 - x_2) \end{bmatrix}$$

$\mathbf{x} = (0, 1)^T$ のとき

$$\mathbf{d} = (2, -20)^T$$

3. (6)

$$\nabla^2 f(\mathbf{x}) \mathbf{d} = -\nabla f(\mathbf{x})$$

$$f(\mathbf{x}) = (x_1 - 1)^2 + 10(x_1^2 - x_2)^2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 1) + 40(x_1^2 - x_2)x_1 \\ -20(x_1^2 - x_2) \end{bmatrix}$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 2 + 120x_1^2 - 40x_2 & -40x_1 \\ -40x_1 & 20 \end{bmatrix}$$

$\mathbf{x} = (0, 0)^T$ のとき

$$\begin{bmatrix} 2 & 0 \\ 0 & 20 \end{bmatrix} \mathbf{d} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \longrightarrow \mathbf{d} = (1, 0)^T$$

表 4.1 関数 (4.19) に対する最急降下法の計算結果

反復 k	$\boldsymbol{x}^{(k)}$	$f(\boldsymbol{x}^{(k)})$	$\ \nabla f(\boldsymbol{x}^{(k)})\ $
0	(0.00000, 1.00000)	0.11000×10^2	0.20099×10^2
1	(0.09988, 0.00115)	0.81098×10^0	0.17737×10^1
2	(0.36070, 0.02723)	0.51452×10^0	0.20677×10^1
3	(0.35167, 0.11761)	0.42069×10^0	0.12174×10^1
4	(0.44424, 0.12687)	0.35853×10^0	0.14166×10^1
5	(0.43824, 0.18689)	0.31583×10^0	0.10381×10^1
10	(0.57217, 0.28642)	0.19981×10^0	0.82338×10^0
20	(0.67723, 0.43291)	0.11080×10^0	0.51716×10^0
30	(0.73949, 0.52795)	0.71432×10^{-1}	0.37970×10^0
40	(0.78289, 0.59811)	0.49328×10^{-1}	0.29770×10^0
50	(0.81555, 0.65307)	0.35472×10^{-1}	0.24213×10^0
100	(0.90619, 0.81570)	0.91002×10^{-2}	0.11011×10^0
200	(0.96896, 0.93721)	0.99136×10^{-2}	0.33934×10^{-1}
300	(0.98869, 0.97691)	0.13148×10^{-3}	0.12106×10^{-1}
400	(0.99575, 0.99130)	0.18520×10^{-4}	0.45109×10^{-2}
500	(0.99838, 0.99669)	0.26650×10^{-5}	0.17065×10^{-2}

表 4.2 関数 (4.19) に対するニュートン法の計算結果

反復 k	$\boldsymbol{x}^{(k)}$	$f(\boldsymbol{x}^{(k)})$	$\ \nabla f(\boldsymbol{x}^{(k)})\ $
0	(0.00000, 0.00000)	0.10000×10^1	0.20000×10^1
1	(0.32341, 0.00000)	0.56717×10^0	0.20919×10^1
2	(0.73455, 0.46247)	0.12990×10^0	0.23209×10^1
3	(0.91297, 0.85632)	0.12775×10^{-1}	0.11054×10^1
4	(1.00450, 1.01041)	0.39429×10^{-4}	0.54177×10^{-1}
5	(0.99997, 0.99995)	0.16624×10^{-8}	0.46482×10^{-3}
6	(1.00000, 1.00000)	0.39340×10^{-17}	0.17062×10^{-7}

表 4.3 関数 (4.19) に対する準ニュートン法 (BFGS 法) の計算結果

反復 k	$\boldsymbol{x}^{(k)}$	$f(\boldsymbol{x}^{(k)})$	$\ \nabla f(\boldsymbol{x}^{(k)})\ $
0	(0.00000, 1.00000)	0.11000×10^2	0.20099×10^2
1	(0.09988, 0.00115)	0.81098×10^0	0.17737×10^1
2	(0.32845, 0.00381)	0.55927×10^0	0.20815×10^1
3	(0.63413, 0.29092)	0.25752×10^0	0.30512×10^1
4	(0.64276, 0.41585)	0.12769×10^0	0.78598×10^0
5	(0.83666, 0.66037)	0.42380×10^{-1}	0.12754×10^1
6	(0.99543, 0.99483)	0.17689×10^{-3}	0.18421×10^0
7	(1.00116, 1.00249)	0.16527×10^{-5}	0.58349×10^{-2}
8	(0.99998, 0.99998)	0.10160×10^{-8}	0.43471×10^{-3}
9	(1.00000, 1.00000)	0.54333×10^{-12}	0.33545×10^{-5}

<最急降下法>

<大域的収束性をもつ> 長所

出発点をどこに選んでも、なんらかの解に収束することが理論的に保証されている

<1次収束> 収束の速さはあまり優れない

<ニュートン法>

<2次収束> 収束が速い

<局所的収束性をもつが、大域的収束性をもたない> 短所

<準ニュートン法>

<超1次収束><ほぼ大域的収束性を保証>

信頼性(大域的収束性)と計算効率(収束の速さ)の両面で非常に優れた方法。実際に広く用いられている。

(a) 最急降下法 (b) ニュートン法 (c) 準ニュートン法

4. (1) (a) (c) (b)

4. (2) (b) (c) (a)

4. (3) (c)

<ナップサック問題>

n 個の品物がある. 品物 i の重さ: a_i kg 利用価値: c_i とする. ナップサックには全部で b kgしか詰めない. 利用価値の総計が最大となる品物を選ぶにはどうすればよいか.

$$\text{目的関数: } \sum_{i=1}^n c_i x_i \quad \rightarrow \text{最大}$$

$$\text{制約条件: } \sum_{i=1}^n a_i x_i \leq b$$

$$x_i = 0, 1 \quad (i=1, \dots, n)$$

$$x_i = \begin{cases} 1, & \text{品物}i \text{を選ぶとき} \\ 0, & \text{品物}i \text{を選ばないとき} \end{cases}$$

例) ナップサック問題

$$\text{目的関数: } \sum_{i=1}^n c_i x_i \quad \longrightarrow \quad \text{最大}$$

$$\text{制約条件: } \sum_{i=1}^n a_i x_i \leq b$$

$$x_i = 0, 1 \quad (i=1, \dots, n)$$

ただし, c_i, a_i, b はすべて正の整数とし,

$$\frac{c_1}{a_1} \geq \frac{c_2}{a_2} \geq \dots \geq \frac{c_n}{a_n}$$

となるように, まえもって変数の添字 i は並べ換えられていると仮定する.

5. (1)

目的関数: $3x_1 + 4x_2 + x_3 + 2x_4 \rightarrow$ 最大

制約条件: $2x_1 + 3x_2 + x_3 + 3x_4 \leq 4$

$$x_i = 0, 1 \quad (i=1, \dots, 4)$$

<連続緩和問題>

目的関数: $3x_1 + 4x_2 + x_3 + 2x_4 \rightarrow$ 最大

制約条件: $2x_1 + 3x_2 + x_3 + 3x_4 \leq 4$

$$0 \leq x_i \leq 1 \quad (i=1, \dots, 4)$$

5. (2)

<連続緩和問題>

目的関数: $3x_1 + 4x_2 + x_3 + 2x_4 \rightarrow$ 最大

制約条件: $2x_1 + 3x_2 + x_3 + 3x_4 \leq 4$

$$0 \leq x_i \leq 1 \quad (i=1, \dots, 4)$$

$x_1 = 1$ 成り立つ

$2 + 3x_2 \leq 4$ より $x_2 = 2/3$

$$x_3 = x_4 = 0$$

最適解: $x = (1, 2/3, 0, 0)^T$ 元の問題の実数最適解

目的関数の値 $z = 17/3$

連続緩和問題は元の問題の目的関数値の上界値を与える。

連続緩和問題の実数最適解により、元の問題の近似最適解を容易に得ることができる。

<分枝図>

- ・最上部の節点: どの変数も固定されていない状態
- ・最下部の節点: すべての変数が0または1に固定された状態
- ・途中の節点: 一部の変数の値が0または1に固定され、それ以外の変数は固定されていないような問題



部分問題

5. (3)

实数最適解： $x = (1, 2/3, 0, 0)^T$

近似最適解： $x = (1, 0, 1, 0)^T$

5. (4)

暫定解： $x = (1, 0, 1, 0)^T$ 暫定値： $z^* = 4$

(a) $x_1=0$ に固定した部分問題

目的関数: $4x_2 + x_3 + 2x_4 \rightarrow$ 最大

制約条件: $3x_2 + x_3 + 3x_4 \leq 4$
 $x_i = 0, 1 \quad (i=2, \dots, 4)$

<連続緩和問題>

目的関数: $4x_2 + x_3 + 2x_4 \rightarrow$ 最大

制約条件: $3x_2 + x_3 + 3x_4 \leq 4$

$0 \leq x_i \leq 1 \quad (i=2, \dots, 4)$

実数最適解: $(x_2, x_3, x_4) = (1, 1, 0)^T$

部分問題の最適解 $x = (0, 1, 1, 0)^T$ が得られた

→ 終端できる

(b) $x_1=1$, $x_2=0$ に固定した部分問題

目的関数: $3 + x_3 + 2x_4$ → 最大

制約条件: $2 + x_3 + 3x_4 \leq 4$

$x_i = 0, 1$ ($i=3, 4$)

<連続緩和問題>

目的関数: $3 + x_3 + 2x_4$ → 最大

制約条件: $2 + x_3 + 3x_4 \leq 4$

$0 \leq x_i \leq 1$ ($i=3, 4$)

実数最適解: $(x_3, x_4) = (1, 1/3)^T$ $z=14/3$

目的関数値が暫定値 $z^*=4$ より大きい

→ 終端できない

(c) $x_1=1$, $x_2=1$ に固定した部分問題

目的関数: $7 + x_3 + 2x_4$ \rightarrow 最大

制約条件: $5 + x_3 + 3x_4 \leq 4$

$x_i = 0, 1$ ($i=3, 4$)

明らかに実行可能解をもたない

\rightarrow 終端できる

5. (5) 分枝限定法の適用

(0) 5. (3)で求めた近似最適解を暫定解とする。

暫定解: $x = (1, 0, 1, 0)^T$ 暫定値: $z^* = 4$

(1) $x_1 = 0$ に固定した部分問題を解く

5. (4)(a)より部分問題の最適解 $x = (0, 1, 1, 0)^T$
が得られたから終端できる

目的関数値 $z = 5$ は暫定値 $z^* = 4$ より大きいから、これを新たな暫定解とし、暫定値 $z^* = 5$ とする。

(2) $x_1=1$ に固定した部分問題を解く

目的関数: $3+4x_2+x_3+2x_4 \longrightarrow$ 最大

制約条件: $2+3x_2+x_3+3x_4 \leq 4$

$x_i=0,1 \quad (i=2,\dots,4)$

<連続緩和問題>

目的関数: $3+4x_2+x_3+2x_4 \longrightarrow$ 最大

制約条件: $2+3x_2+x_3+3x_4 \leq 4$

$0 \leq x_i \leq 1 \quad (i=2,\dots,4)$

実数最適解: $(x_2, x_3, x_4) = (2/3, 0, 0)^T$

目的関数値 $z=17/3$ は暫定値 $z^*=5$ より大きいから、最適解を含む可能性がある。そこで新たな部分問題を生成する。

(3) $x_1=1, x_2=0$ に固定した部分問題を解く

5. (4)(b)より部分問題の実数最適解: $(x_3, x_4) = (1, 1/3)^T$ $z=14/3$ が得られる。しかし $z=14/3$ は暫定値 $z^*=5$ より小さい。
したがって、この部分問題は終端する。

(4) $x_1=1, x_2=1$ に固定した部分問題を解く

5. (4)(c)よりこの部分問題は実行可能解をもたない。したがって、この部分問題は終端する。

(5) すべての部分問題が終端したので、暫定解 $x = (0, 1, 1, 0)^T$ が最適解となる。

6. (1)

現在の解(基底解)に対して、非基底変数と基底変数を1つずつ入れ替えることにより生成される解(基底解)の集合

6. (2)

現在の解(フロー)に対する残余ネットワークにおけるいずれかのパスにフローを流すことにより得られる解(フロー)の集合

6. (3)

現在の解(点)における勾配の逆方向にある解(点)の集合

6. (4)

現在の解(巡回路)に対して2本の枝を付け替えることにより得られる解(巡回路)の集合

7 <メタヒューリスティクス>

改悪となるような解への移動も許すことによって、好ましくない局所最適解に補足されることを避けようとする方法

<焼きなまし法(シミュレーテッド・アニーリング法)>

<特徴>

改悪となる解を採用する確率を温度と呼ばれるパラメータを用いて変化させる

<タブー探索法>

過去の探索で現れた解や移動のパターンをタブーリストと呼ばれる集合の形で記憶しておき、そのリストに含まれない解のなかで最良のものに移動する