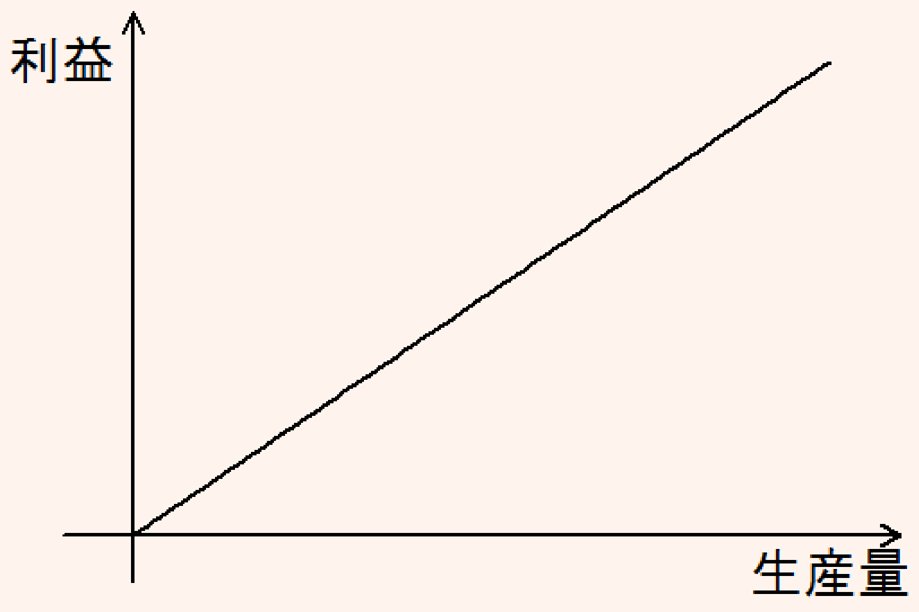
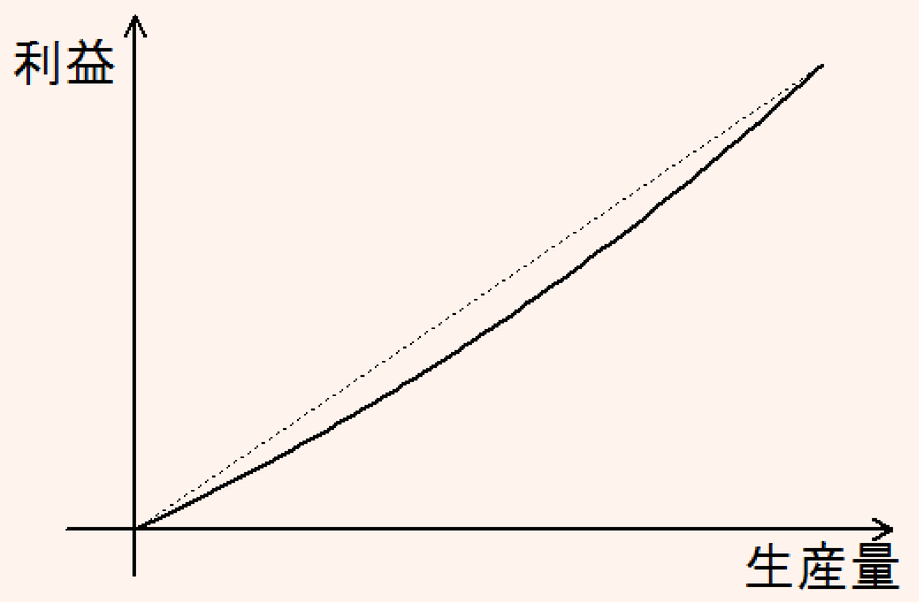
非線形最適化問題

◆生産量と利益見込み(線形関係)



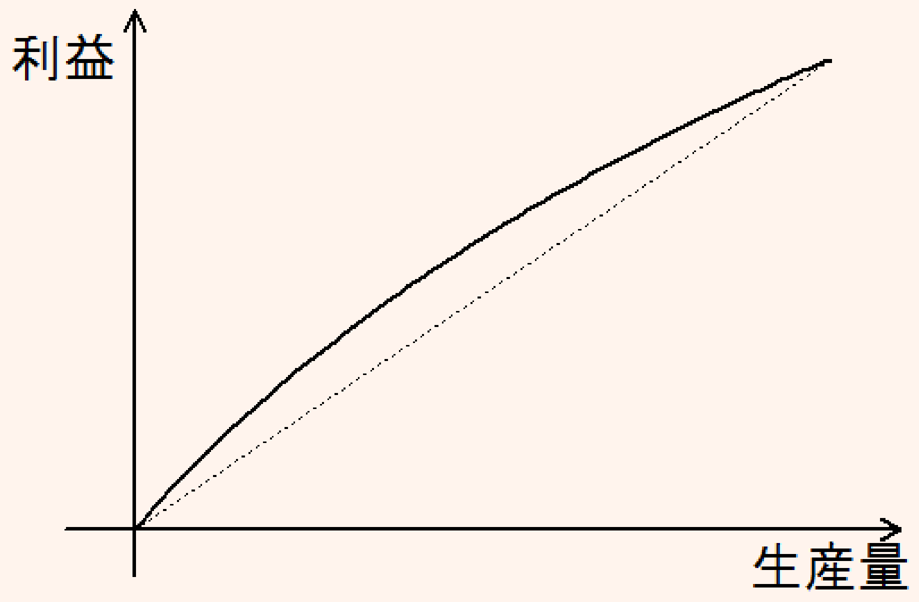
非線形最適化問題

◆生産量と利益見込み(非線形関係)



非線形最適化問題

◆生産量と利益見込み(非線形関係)

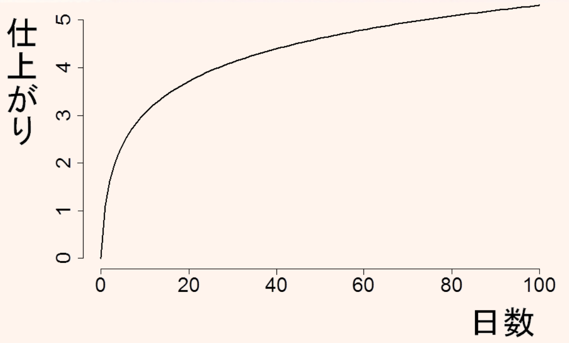


非線形最適化問題

◆制約のある非線形最適化問題

- 工程 $i\ (i = 1, 2, \cdots, n)$ に x(日)かけた時の仕上がりの程度

$$\log(a_i x + 1) \qquad a_i > 0$$



- 各工程は最低 $t_i (> 0)$ 日かかる
- 全工程を T 日以内に終わらせなければならない
- 仕上がりの和を最大にするには，各工程に何日かければよいか

非線形最適化問題

◆制約のある非線形最適化問題

- 工程 i に x_i 日かける
- 目的関数 z は各工程の仕上がり程度の総和

$$z = \sum_{i=1}^{n} \log(a_i x_i + 1)$$

非線形最適化問題

◆制約のある非線形最適化問題

- 工程iには最低t_i(日)かかる

$$x_i \geq t_i \ \ (i = 1, 2, \cdots n)$$

- 全工程をT日以内に終わらせる

$$\sum_{i=1}^{n} x_i \leq T$$

非線形最適化問題

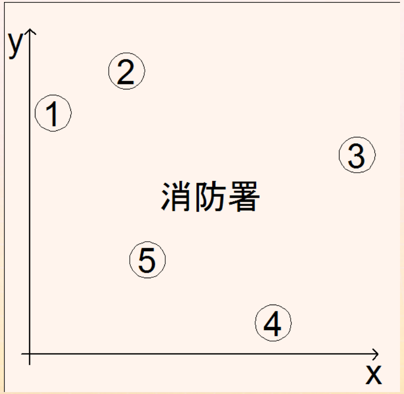
◆制約のある非線形最適化問題

$$
\begin{aligned}
\text{最大化}\quad & z = \sum_{i=1}^{n} \log(a_i x_i + 1) \\
\text{制約条件}\quad & x_i \geq t_i \ \ (i = 1, 2, \cdots, n) \\
& \sum_{i=1}^{n} x_i \leq T
\end{aligned}
$$

非線形最適化問題

◆制約のない非線形最適化問題

- 5つの人口密集地区 (1, 2, 3, 4, 5)
- 消防署はこれら5つの地区への直線距離の和が最小になる位置に設置
- 消防署を設置すべき位置を求めよ



地区	x	y
1	1	12
2	4	14
3	15	10
4	11	2
5	5	5

非線形最適化問題

◆制約のない非線形最適化問題

- 消防署の位置の座標を (x, y)
- 地区 i $(i = 1, 2, 3, 4, 5)$ の位置の座標　(x_i, y_i)

- 消防署と地区 i の直線距離は,

$$\sqrt{(x - x_i)^2 + (y - y_i)^2}$$

非線形最適化問題

◆制約のない非線形最適化問題

- 消防署の位置の座標を (x, y)
- 地区 i $(i = 1, 2, 3, 4, 5)$ の位置の座標を (x_i, y_i)
- 目的関数 z は消防署と各地区の直線距離の和で

$$z = \sum_{i=1}^{5} \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

- z を最小化する x, y を求める問題

非線形最適化問題

◆非線形最適化問題の一般形

$$
\begin{aligned}
\text{最小化} \quad & f(\boldsymbol{x}) \\
\text{制約} \quad & g_i(\boldsymbol{x}) = 0 \qquad (i = 1, 2, \cdots, m) \\
& h_j(\boldsymbol{x}) \leq 0 \qquad (j = 1, 2, \cdots, l)
\end{aligned}
$$

- $\boldsymbol{x}$はn変数のベクトル$[x_1, x_2, \cdots, x_n]^T$

非線形最適化問題

◆今回の前提

- 制約のない非線形最適化問題の最適化
- 問題は最小化問題
- 目的関数 $f(\boldsymbol{x})$ は考慮すべき$\boldsymbol{x}$の領域において2階微分可能
- 2階微分は連続関数

非線形最適化問題

◆最適性の条件

- 任意の $\boldsymbol{x}$ に対して $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ を満たす $\boldsymbol{x}^*$
  $\cdots$ 関数 f の最小化問題の大域最適解
- $\boldsymbol{x}^*$ の周囲の $\boldsymbol{x}$ に対して $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ を満たす $\boldsymbol{x}^*$
  $\cdots$ 関数 f の最小化問題の局所最適解
- 大域最適解は局所最適解であるが，局所最適解は一般には大域最適解とはならない

非線形最適化問題

- 1変数関数 $f(x)$
- 関数 $f(x)$ の増減を調べるには，関数の変化率を表す(1階)導関数(微分)

$$f'(x) = \frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

- 導関数の変化率を表す2階導関数

$$f''(x) = \frac{d^2 f(x)}{dx^2} = \lim_{\Delta x \to 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}$$

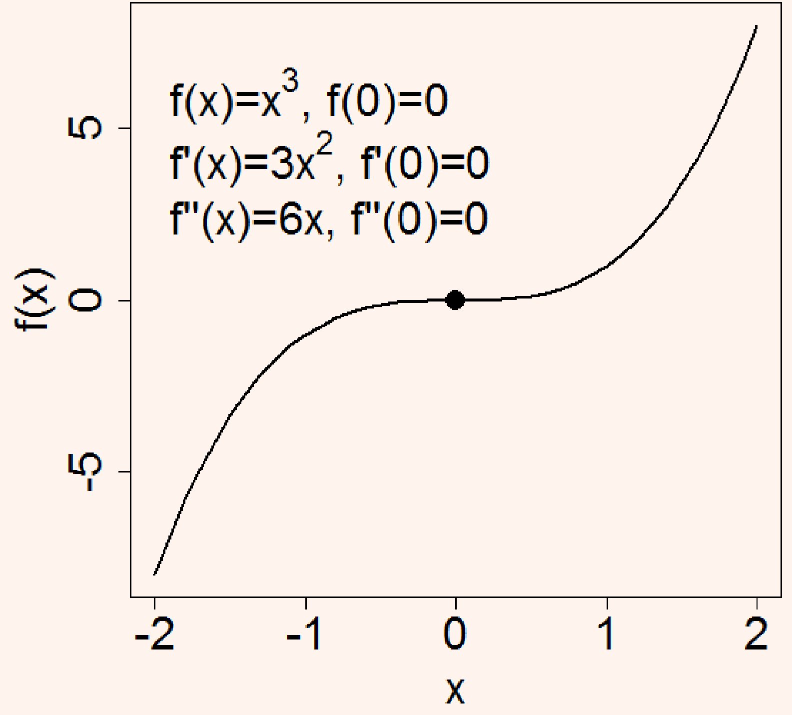
非線形最適化問題

◆最適性の条件

- $f'(x)$ は x における $f(x)$ の変化率 (接線の傾き)
- x において，$f'(x)$ が正なら x において $f(x)$ は増加
- $f'(x)$ が負なら x において $f(x)$ は減少
- x において $f(x)$ が極小値あるいは極大値
  $\rightarrow$ x において $f'(x) = 0$
- $f'(x) = 0$ なら必ずしも極小値あるいは極大値になっている訳ではない

非線形最適化問題

◆最適性の条件



非線形最適化問題

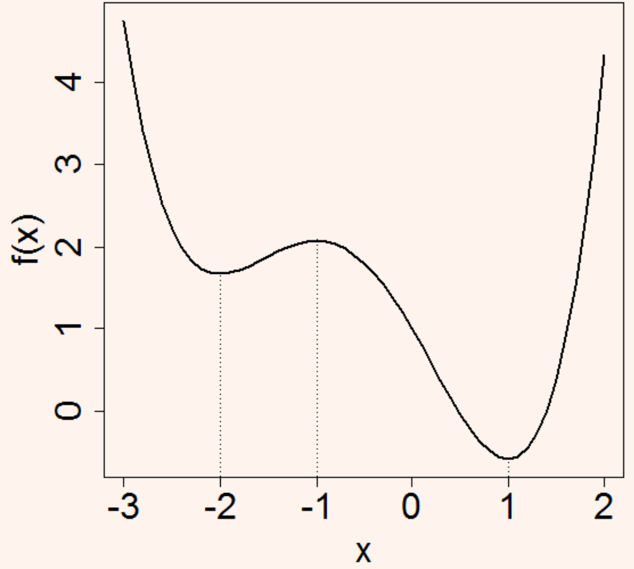
◆最適性の条件

- $f'(x)=0$のとき，xにおいて$f(x)$が極小値であるか極大値であるかの判別
- $f''(x)$が正なら$f(x)$は下に凸な曲線$(\smile)$ $\cdots$ 極小値
- $f''(x)$が負なら$f(x)$は上に凸な曲線$(\frown)$ $\cdots$ 極大値
- $f''(x)$が0の場合，これだけからでは最適性の判定はできない

非線形最適化問題

◆最適性の条件

$$f(x) = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{2}x^2 - 2x + 1$$



非線形最適化問題

◆最適性の条件

$$f(x) = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{2}x^2 - 2x + 1$$

$$f'(x) = x^3 + 2x^2 - x - 2 = (x+2)(x+1)(x-1)$$

- $x = -2,\ -1,\ 1 \ \rightarrow\ f'(x) = 0$

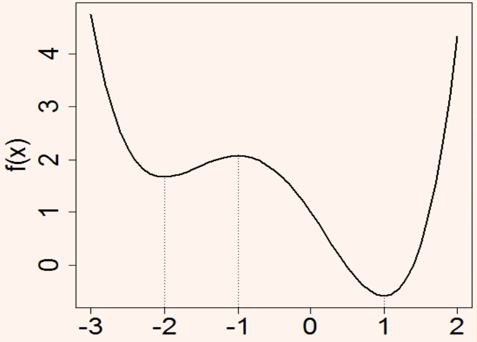
$$f''(x) = 3x^2 + 4x - 1$$

- $x = \dfrac{-2 \pm \sqrt{7}}{3} \ \rightarrow\ f''(x) = 0$

非線形最適化問題

◆最適性の条件

$$f(x) = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{2}x^2 - 2x + 1$$

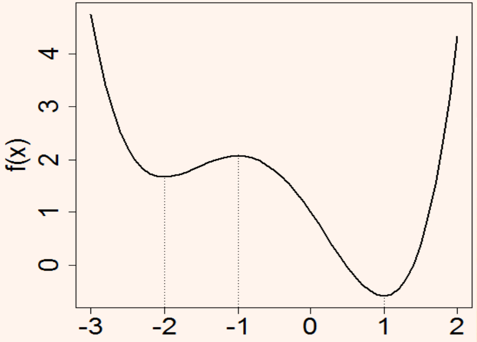


x	< -2	-2		$\frac{-2-\sqrt{7}}{3}$		-1		$\frac{-2+\sqrt{7}}{3}$		1	> 1
$f'(x)$	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$
$f''(x)$	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	$+$	$+$
$f(x)$	$\smile$	$\frac{5}{3}$ 極小	$\smile$		$\frown$	$\frac{25}{12}$ 極大	$\frown$		$\smile$	$\frac{-7}{12}$ 極小	$\smile$

非線形最適化問題

◆最適性の条件

$$f(x) = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{2}x^2 - 2x + 1$$

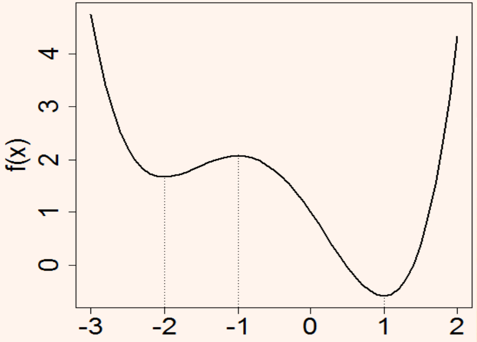


x	< -2	-2		$\frac{-2-\sqrt{7}}{3}$		-1		$\frac{-2+\sqrt{7}}{3}$		1	> 1
$f'(x)$	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$
$f''(x)$	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	$+$	$+$
$f(x)$	$\smile$	$\frac{5}{3}$ 極小	$\smile$		$\frown$	$\frac{25}{12}$ 極大	$\frown$		$\smile$	$\frac{-7}{12}$ 極小	$\smile$

非線形最適化問題

◆最適性の条件

$$f(x) = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{2}x^2 - 2x + 1$$



x	< -2	-2		$\frac{-2-\sqrt{7}}{3}$		-1		$\frac{-2+\sqrt{7}}{3}$		1	> 1
$f'(x)$	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$
$f''(x)$	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	$+$	$+$
$f(x)$	$\smile$	$\frac{5}{3}$ 極小	$\smile$		$\frown$	$\frac{25}{12}$ 極大	$\frown$		$\smile$	$\frac{-7}{12}$ 最小	$\smile$

非線形最適化問題

◆最適性の条件

- n 変数関数 $f(\boldsymbol{x})$ $(\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T)$ の偏導関数

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, \cdots, x_i + \Delta x_i, x_{i+1}, \cdots, x_n) - f(x_1, \cdots, x_i, x_{i+1}, \cdots, x_n)}{\Delta x_i}$$

$$\frac{f(x_1, \cdots, x_i + \Delta x_i, x_{i+1}, \cdots, x_n) - f(x_1, \cdots, x_i, x_{i+1}, \cdots, x_n)}{\Delta x_i}$$

- x_i のみを独立変数と見なして微分

非線形最適化問題

◆最適性の条件

- 偏導関数 $\frac{\partial f(\boldsymbol{x})}{\partial x_i}$,　$f_{x_i}(\boldsymbol{x})$ を成分とする n 次元ベクトル

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

勾配ベクトル

非線形最適化問題

◆最適性の条件

- 2階の偏導関数$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}$, $f_{x_i x_j}(\boldsymbol{x})$を成分とする$n \times n$行列

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

ヘッセ行列

非線形最適化問題

◆最適性の条件

$n \times n$行列$\boldsymbol{M}$

- 任意のn次元ベクトル$\boldsymbol{x}$に対して

$$\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} \geq 0$$

を満たすとき，行列$\boldsymbol{M}$は半正定値

- 任意の実数xに対して

$$mx^2 \geq 0$$

を満たすとき，$m \geq 0$

非線形最適化問題

◆最適性の条件

$n \times n$行列$\boldsymbol{M}$

- 零ベクトルを除く任意のn次元ベクトル$\boldsymbol{x}$に対して

$$\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} > 0$$

を満たすとき，行列$\boldsymbol{M}$は正定値

- 0を除く任意の実数xに対して

$$mx^2 > 0$$

を満たすとき，$m > 0$

非線形最適化問題

◆最適性の条件

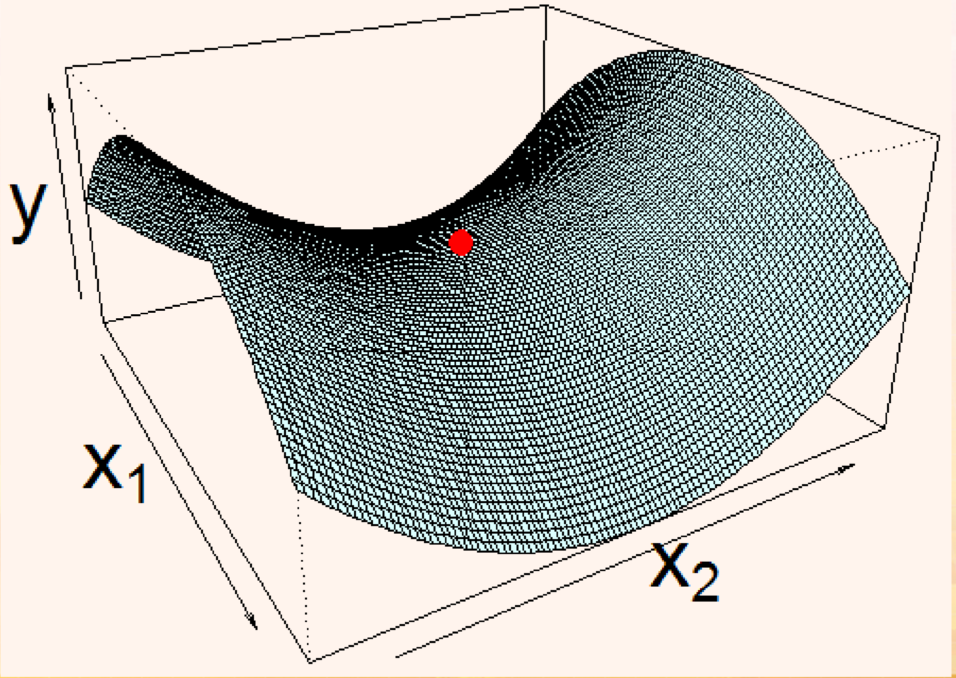
- $\boldsymbol{x}$が局所最適解であれば

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

- $\boldsymbol{x}$が局所最適解であるための必要条件
  $\cdots$ 1次の必要条件
- $\nabla f(\boldsymbol{x}) = \mathbf{0}$を満たす$\boldsymbol{x}$
  $\cdots$ $f(\boldsymbol{x})$の停留点

非線形最適化問題

◆最適性の条件



非線形最適化問題

◆最適性の条件

- $\boldsymbol{x}$ が局所最適解であれば

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

かつ

$$\nabla^2 f(\boldsymbol{x})\text{は半正定値}$$

- $\boldsymbol{x}$が局所最適解であるための必要条件
… 2次の必要条件

非線形最適化問題

◆最適性の条件

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

かつ

$$\nabla^2 f(\boldsymbol{x})\text{は正定値}$$

であれば

- $\boldsymbol{x}$ は局所最適解
- $\boldsymbol{x}$が局所最適解であるための十分条件
  … 2次の十分条件

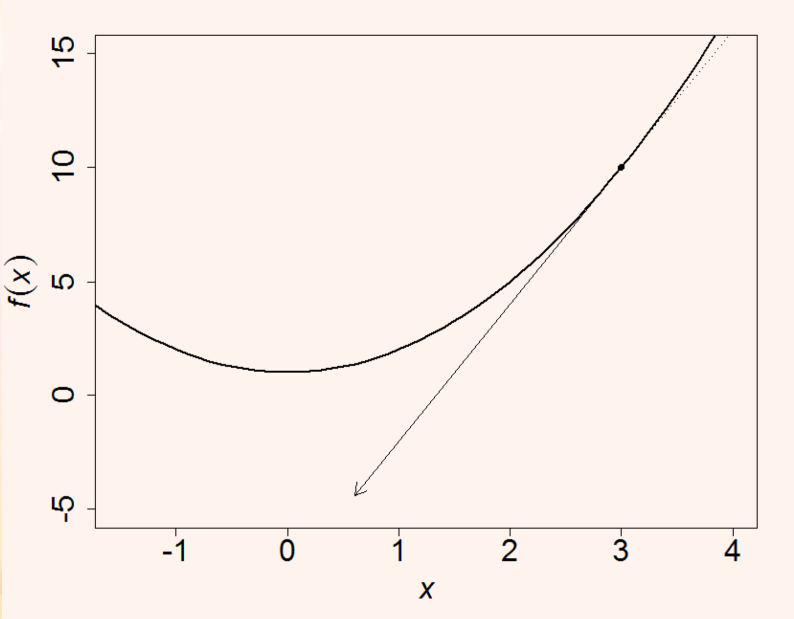
最急降下法

最急降下法

- $f'(x)$ … 関数 $f(x)$ の増加の割合
- x_0 において

$$f'(x_0) > 0 \quad x \text{を減少}$$
$$f'(x_0) < 0 \quad x \text{を増加}$$

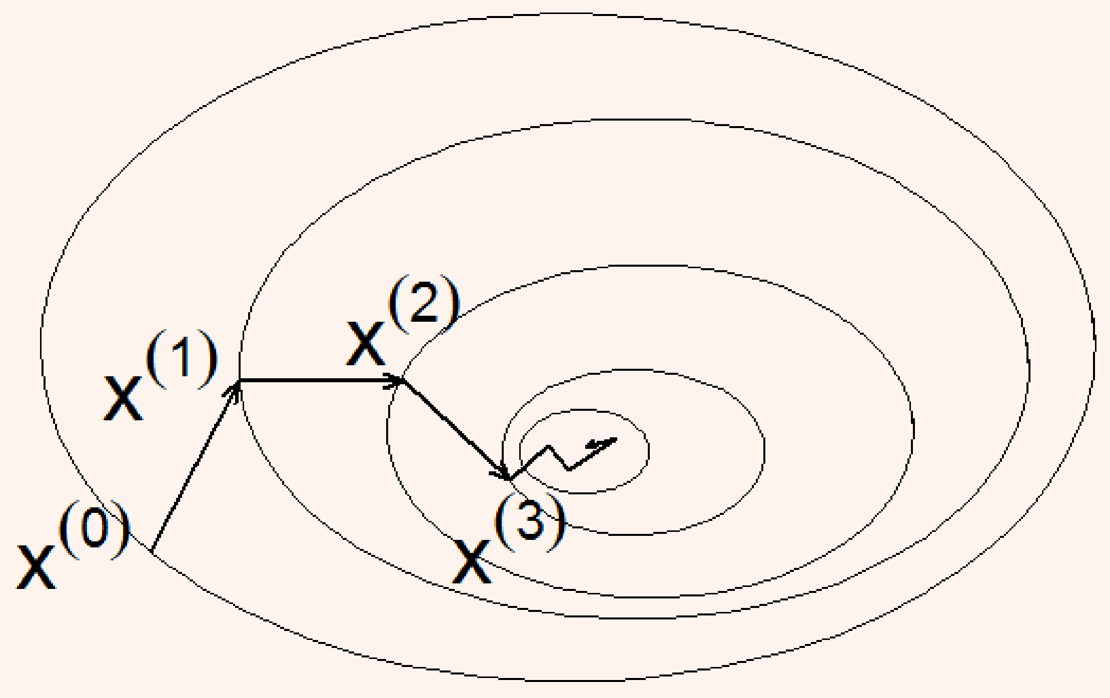
… $f(x)$ を減少



最急降下法

- 勾配ベクトル $\nabla f(\boldsymbol{x})$ の向き
  … 関数 $f(\boldsymbol{x})$ を最も増加させる方向
- 勾配ベクトルの反対方向に $\boldsymbol{x}$ を変化させる
  … $f(\boldsymbol{x})$ を減少
- これを繰り返すことにより局所最適解に到達
  … 最急降下法

最急降下法



最急降下法

- k 回目の繰り返しにおいて $\boldsymbol{x}^{(k)}$ であるとき

$$\boldsymbol{x}^{(k+1)} \leftarrow \boldsymbol{x}^{(k)} - \alpha^{(k)} \nabla f(\boldsymbol{x}^{(k)})$$

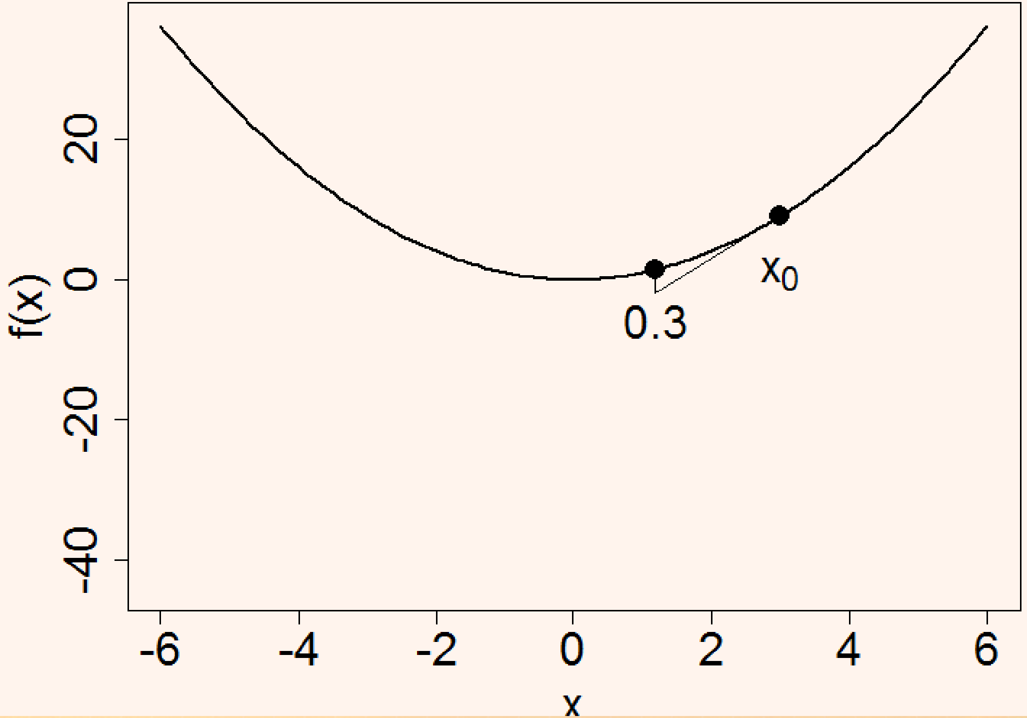
  と更新

- $\alpha^{(k)}$ は正の数

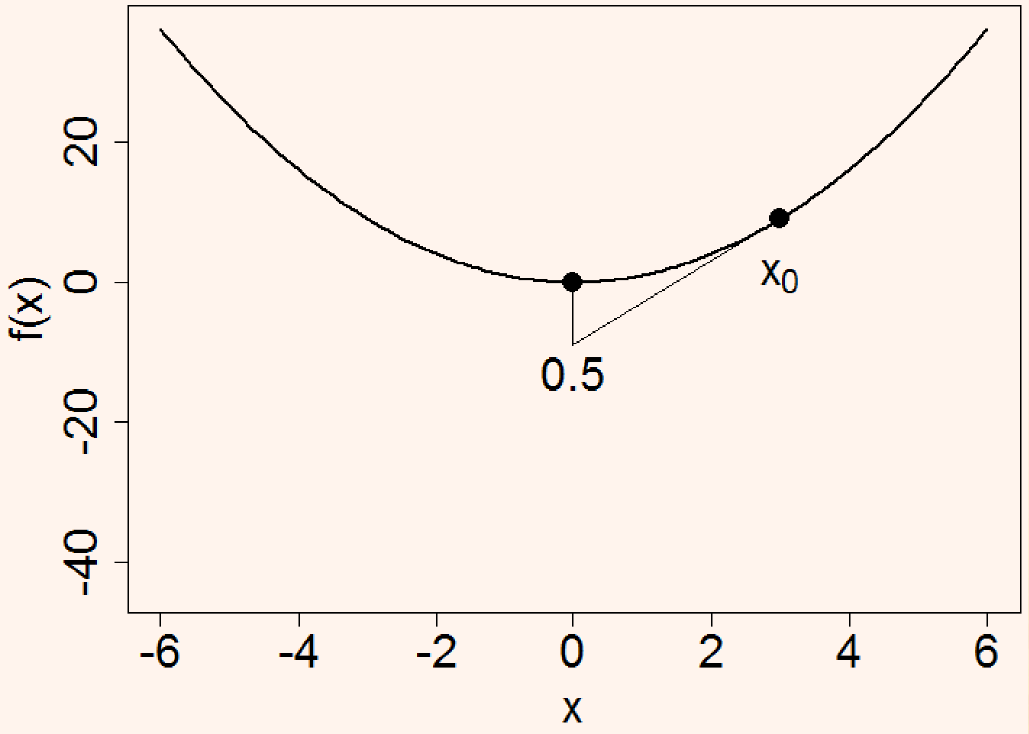
$$f(\boldsymbol{x}^{(k)} - \alpha^{(k)} \nabla f(\boldsymbol{x}^{(k)}))$$

  が極小に(近く)なるように定める

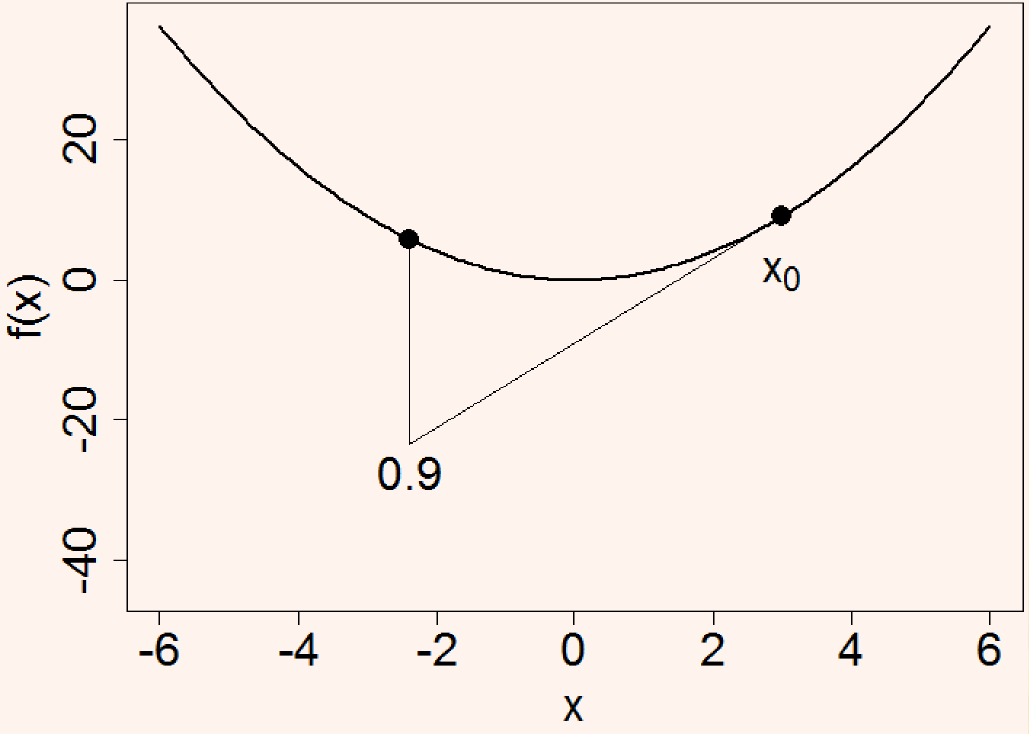
最急降下法



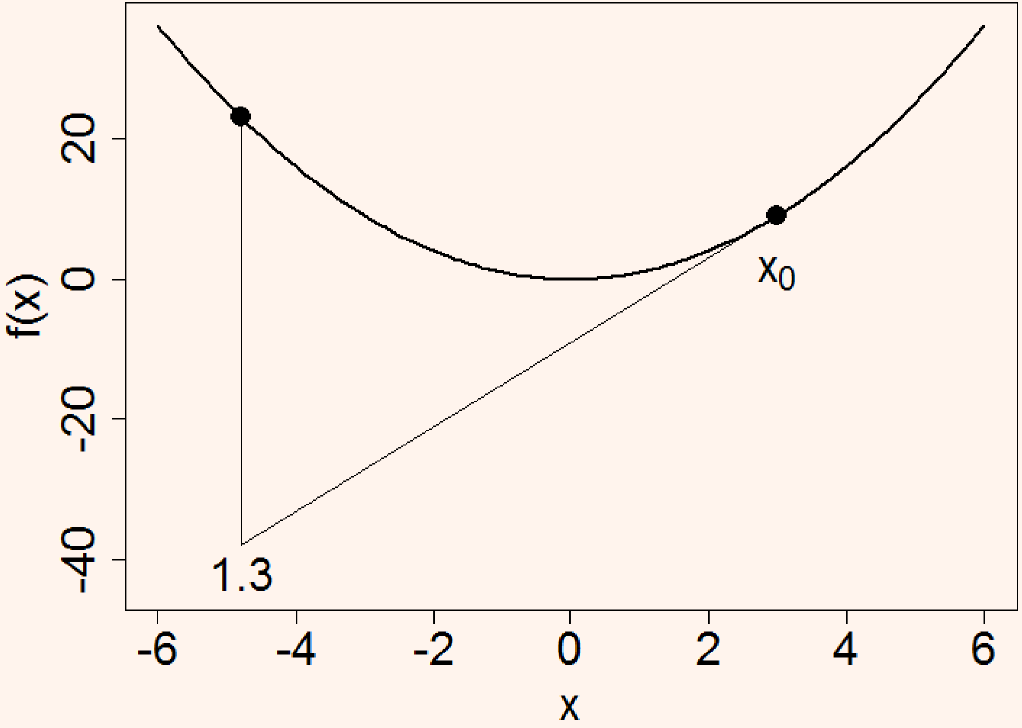
最急降下法



最急降下法



最急降下法



最急降下法

(0) $\boldsymbol{x}$の適当な初期値$\boldsymbol{x}^{(0)}$を定める　$k \leftarrow 0$

(1) $\nabla f(\boldsymbol{x}^{(k)}) = \boldsymbol{0}$ならば$\boldsymbol{x}^{(k)}$を局所最適解として終了
そうでなければ(2)へ

(2) $f(\boldsymbol{x}^{(k)} - \alpha^{(k)}\nabla f(\boldsymbol{x}^{(k)}))$が極小に(近く)なるような$\alpha^{(k)}$を求め

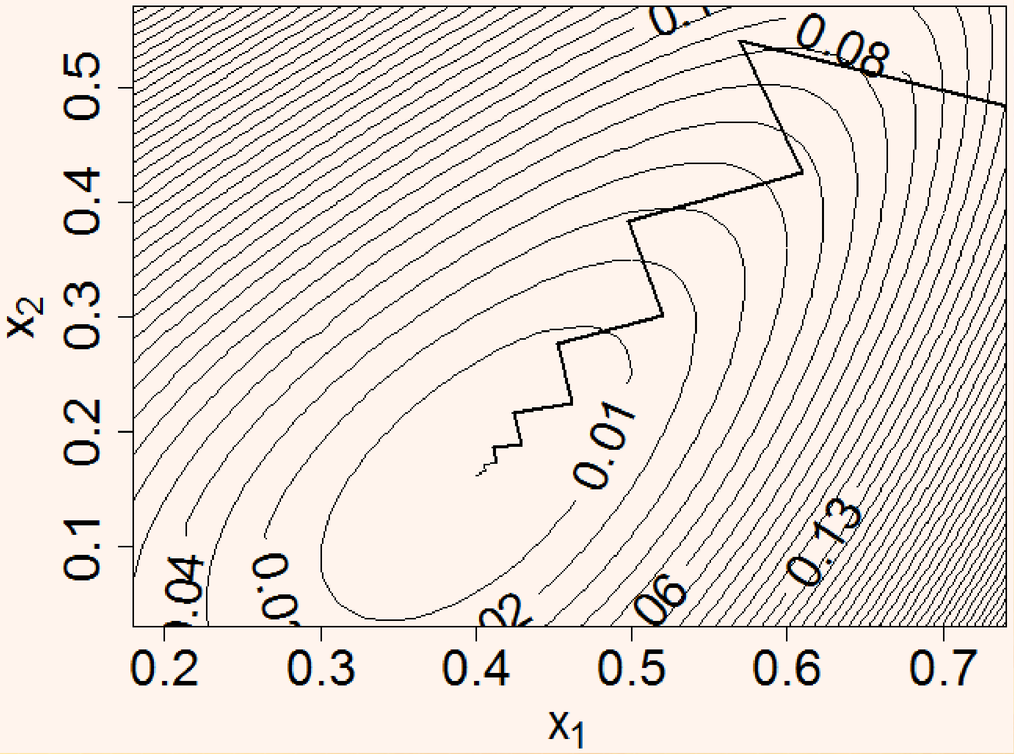
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha^{(k)}\nabla f(\boldsymbol{x}^{(k)})$$

により$\boldsymbol{x}^{(k)}$を更新　$k \leftarrow k+1$として(1)へ

最急降下法

- 実際の終了基準
  小さな正の数εに対して
  $|\nabla f(\boldsymbol{x}^{(k)})| < \varepsilon$
- $\boldsymbol{x}$の任意の初期値に対して局所最適解に収束する
  … 大域収束性
- 収束までに多数の繰り返し数を要することが多い

最急降下法



ニュートン法

ニュートン法

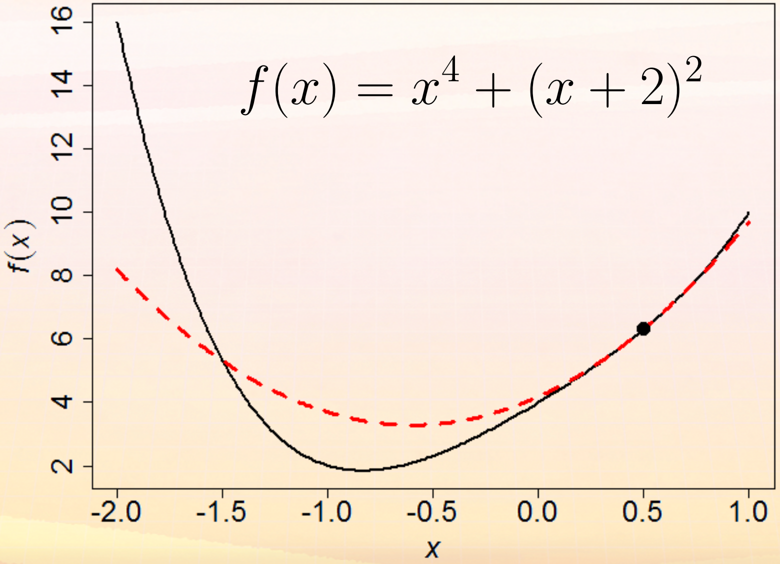
- 1変数関数$f(x)$は，定数$x^{(k)}$のまわりでテイラー展開により

$$
\begin{aligned}
f(x) =& f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2!}f''(x^{(k)}) \left(x - x^{(k)}\right)^2 \\
&+ \frac{1}{3!}f'''(x^{(k)}) \left(x - x^{(k)}\right)^3 + \cdots
\end{aligned}
$$

と近似できる

- 1次関数，2次関数で近似することが多い

ニュートン法



$$g(x) = f(0.5) + f'(0.5)(x-0.5) + \frac{1}{2}f''(0.5)(x-0.5)^2$$

ニュートン法

- $f(x)$ を $x^{(k)}$ のまわりで2次関数$g(x)$ と近似する

$$\begin{aligned} g(x) = \ & f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) \\ & + \frac{1}{2} f''(x^{(k)})(x - x^{(k)})^2 \end{aligned}$$

- $f''(x^{(k)}) > 0$ と仮定すれば
  最適性の1次の必要条件

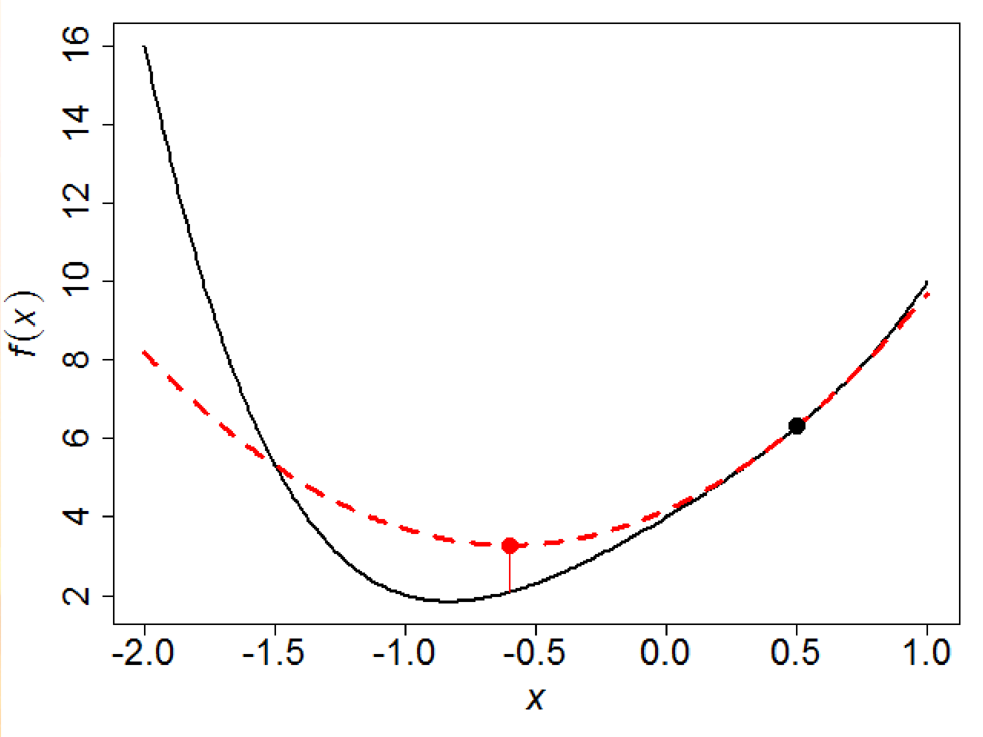
$$g'(x) = 0$$

  を満たす x において $g(x)$ は最小

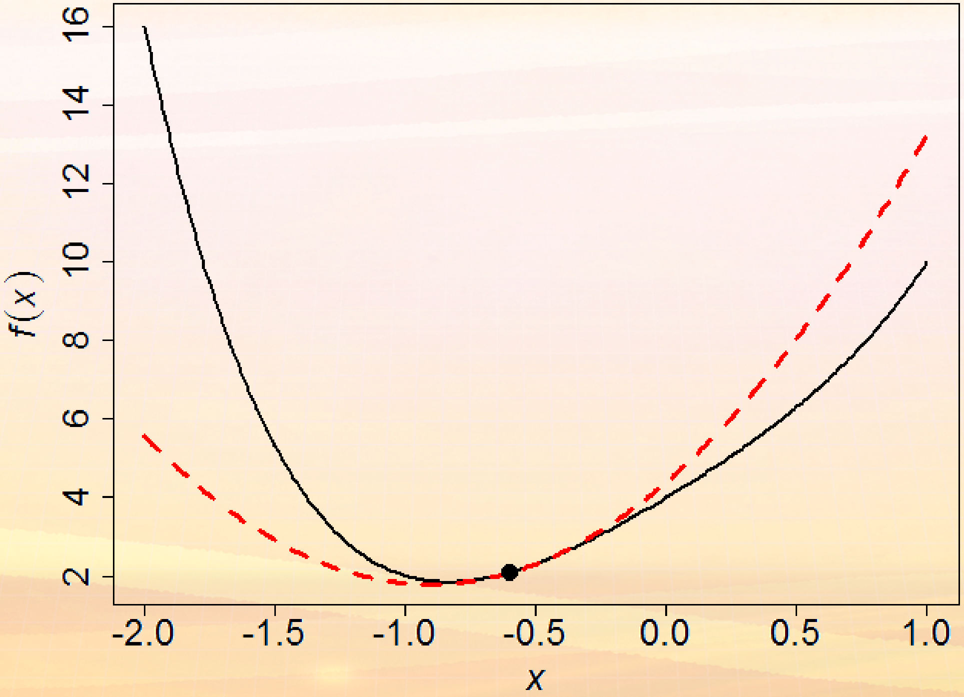
ニュートン法

- 最適解 x は
$$g'(x) = f'(x^{(k)}) + f''(x^{(k)})(x - x^{(k)}) = 0$$
を解いて
$$x = x^{(k)} - f'(x^{(k)})/f''(x^{(k)})$$
- x は $f(x)$ の近似式から導いたので近似解
- 更新を繰り返して $f(x)$ の局所最適解に到達
… ニュートン法

ニュートン法



ニュートン法



ニュートン法

- 多変数関数 $f(\boldsymbol{x})$ は，定数ベクトル $\boldsymbol{x}^{(k)}$ のまわりでテイラー展開により

$$\begin{aligned} f(\boldsymbol{x}) \simeq g(\boldsymbol{x}) = \ & f(\boldsymbol{x}^{(k)}) + \nabla f(\boldsymbol{x}^{(k)})^T(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \\ & + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \\ & + \cdots \end{aligned}$$

と近似できる

ニュートン法

- $f(\boldsymbol{x})$ を $\boldsymbol{x}^{(k)}$ のまわりで2次関数 $g(\boldsymbol{x})$ と近似する

$$g(\boldsymbol{x}) = f(\boldsymbol{x}^{(k)}) + \nabla f(\boldsymbol{x}^{(k)})^T(\boldsymbol{x} - \boldsymbol{x}^{(k)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)})$$

- ヘッセ行列 $\nabla^2 f(\boldsymbol{x}^{(k)})$ が正定値と仮定すれば最適性の1次の必要条件

$$\nabla g(\boldsymbol{x}) = \mathbf{0}$$

を満たす $\boldsymbol{x}$ において $g(\boldsymbol{x})$ は最小

ニュートン法

- 最適解$\boldsymbol{x}$は
$$\nabla g(\boldsymbol{x}) = \nabla f(\boldsymbol{x}^{(k)}) + \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) = \boldsymbol{0}$$
を解いて
$$\boldsymbol{x} = \boldsymbol{x}^{(k)} - \nabla^2 f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})$$
- $\boldsymbol{x}$は$f(\boldsymbol{x})$の近似式から導いたので近似解
- 更新を繰り返して$f(\boldsymbol{x})$の局所最適解に到達
… ニュートン法

ニュートン法

(0) $\boldsymbol{x}$の適当な初期値$\boldsymbol{x}^{(0)}$を定める　$k \leftarrow 0$

(1) $\nabla f(\boldsymbol{x}^{(k)}) = \boldsymbol{0}$ならば
$\boldsymbol{x}^{(k)}$を局所最適解として計算終了
そうでなければ(2)へ

(2) $\boldsymbol{x}^{(k+1)} \leftarrow \boldsymbol{x}^{(k)} - \nabla^2 f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})$
により$\boldsymbol{x}^{(k)}$を更新　$k \leftarrow k+1$として，(1)へ

ニュートン法

◆数値例

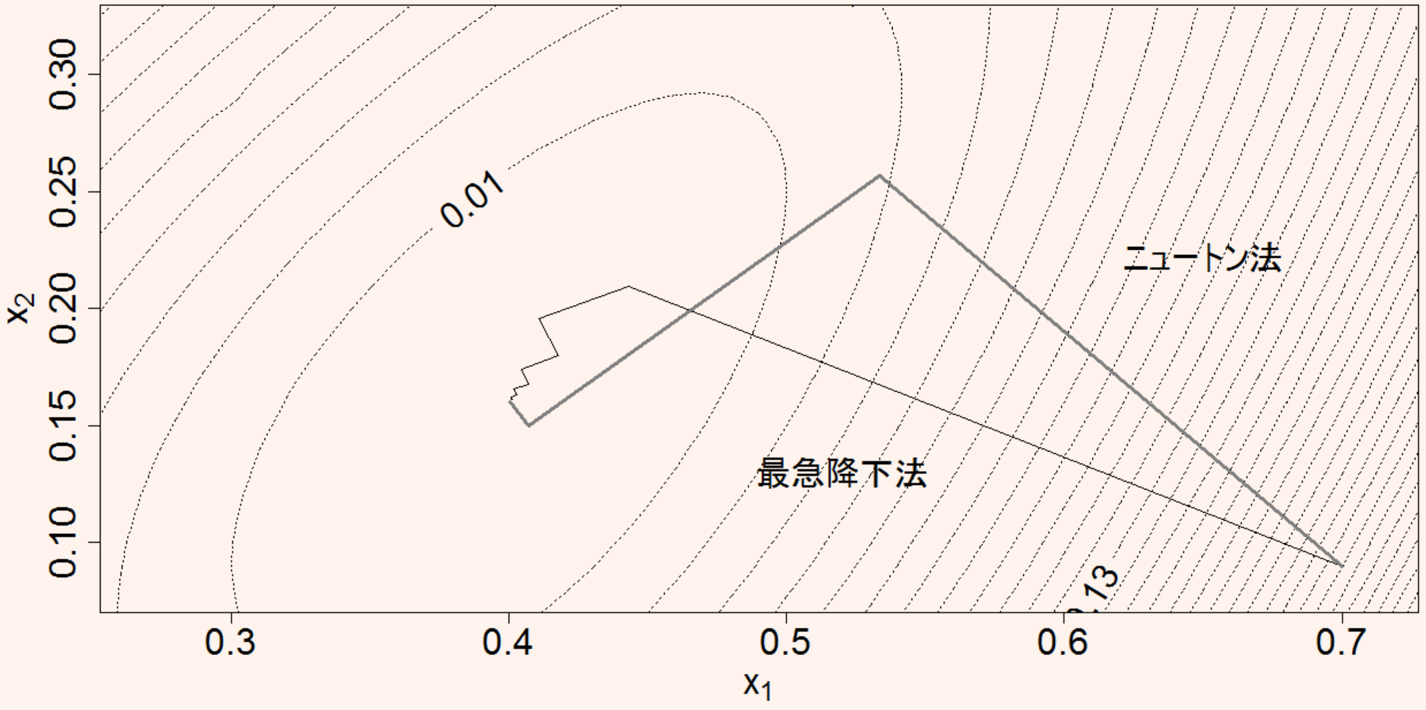
$$f(\boldsymbol{x}) = (x_1 - 0.4)^2 + (x_1^2 - x_2)^2$$

を最小化する$\boldsymbol{x}$を最急降下法とニュートン法を用いて求める

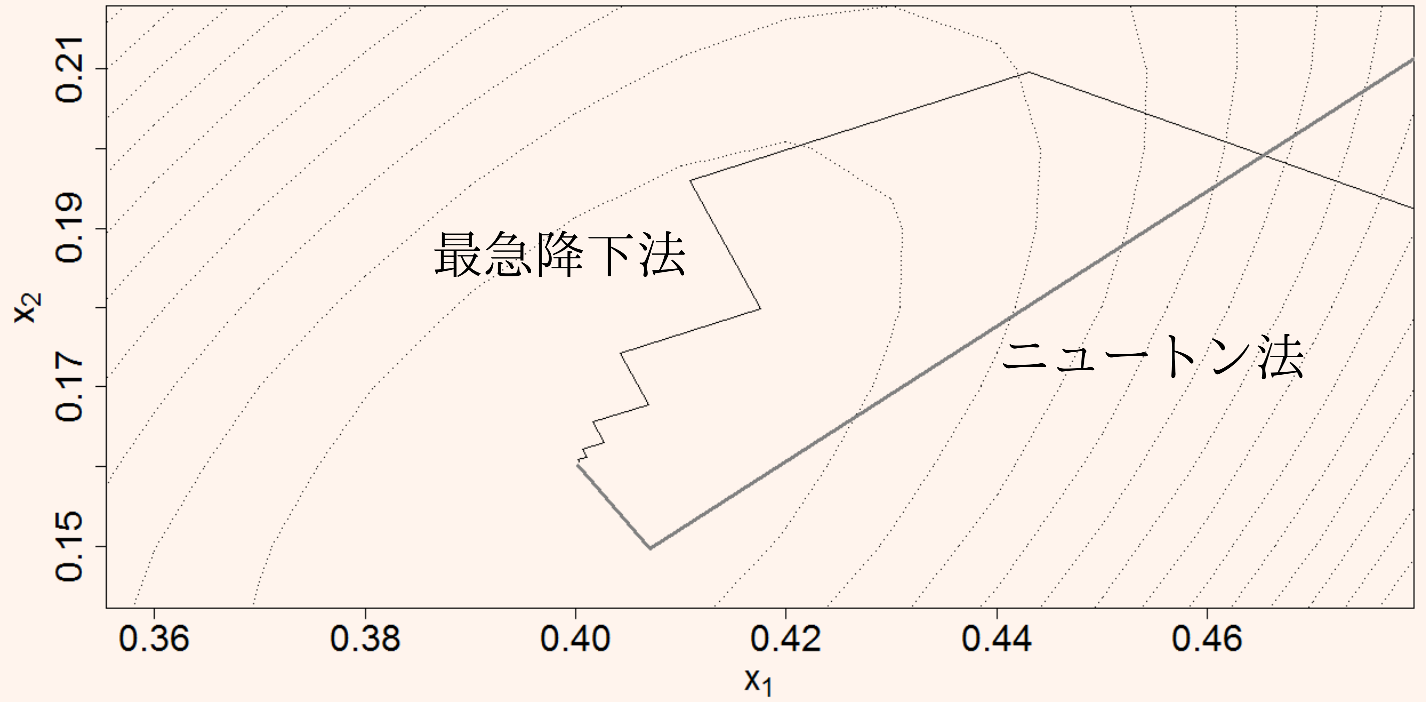
$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 2(x_1 - 0.4) + 4x_1(x_1^2 - x_2) \\ -2(x_1^2 - x_2) \end{bmatrix},$$

$$\nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} 12x_1^2 - 4x_2 + 2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix}$$

ニュートン法



ニュートン法



ニュートン法

- 局所最適解への収束が極めて速い
- ヘッセ行列が常に正定値であることを仮定
  これが満たされない場合
  → 局所最適解への収束が保証されない
- 初期値が局所最適解に十分近ければ,
  局所最適解への収束が保証される… 局所収束性

ニュートン法

◆準ニュートン法

- ヘッセ行列を勾配ベクトルなどから逐次近似

BFGS法

$$\boldsymbol{B}^{(k+1)} = \boldsymbol{B}^{(k)} + \frac{\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)^T}}{\boldsymbol{y}^{(k)^T}\boldsymbol{s}^{(k)}} - \frac{\boldsymbol{B}^{(k)}\boldsymbol{s}^{(k)}\boldsymbol{s}^{(k)^T}\boldsymbol{B}^{(k)^T}}{\boldsymbol{s}^{(k)^T}\boldsymbol{B}^{(k)}\boldsymbol{s}^{(k)}}$$

$$\boldsymbol{B}^{(0)} = \boldsymbol{I},\ \boldsymbol{s}^{(k)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)},\ \boldsymbol{y}^{(k)} = \nabla f(\boldsymbol{x}^{(k+1)}) - \nabla f(\boldsymbol{x}^{(k)})$$