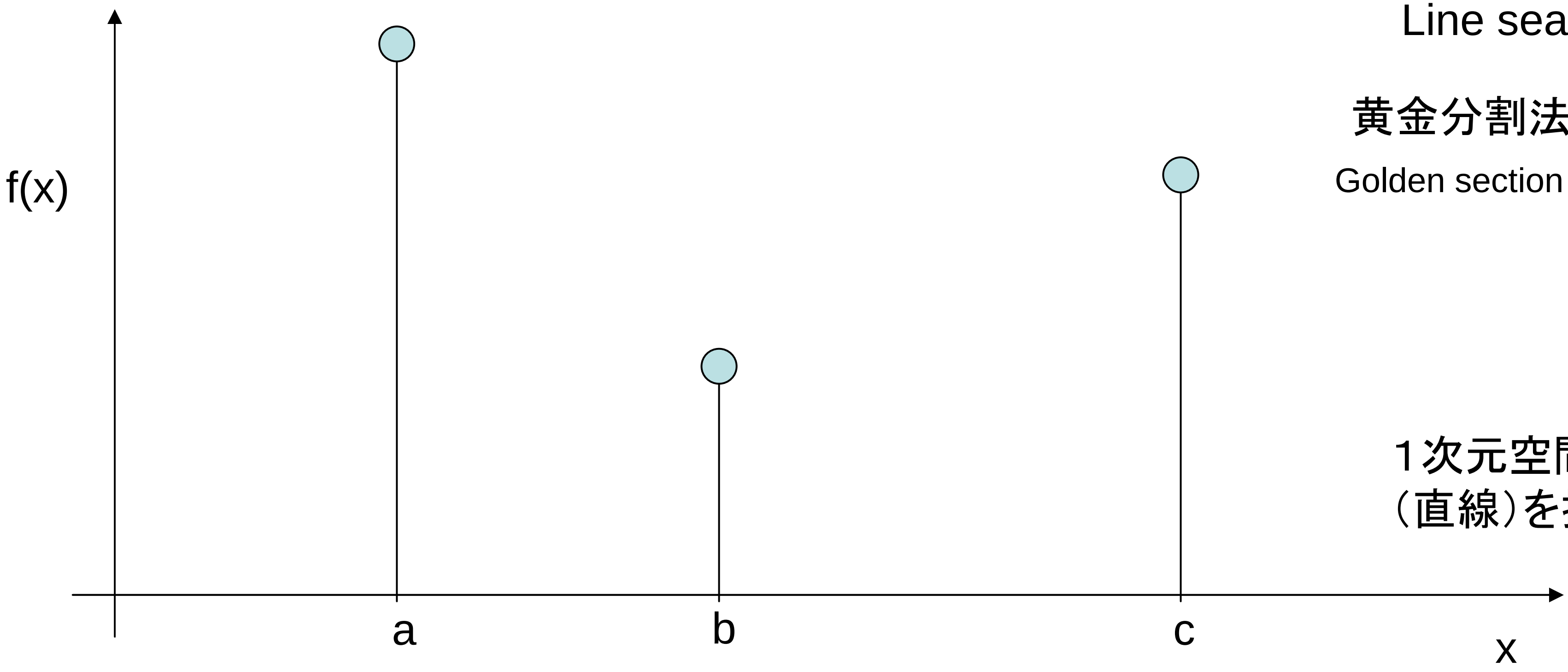
【制約のない関数最適化】1次元の最小化：ラインサーチ

Line search



黄金分割法

Golden section search

1次元空間上
（直線）を探索

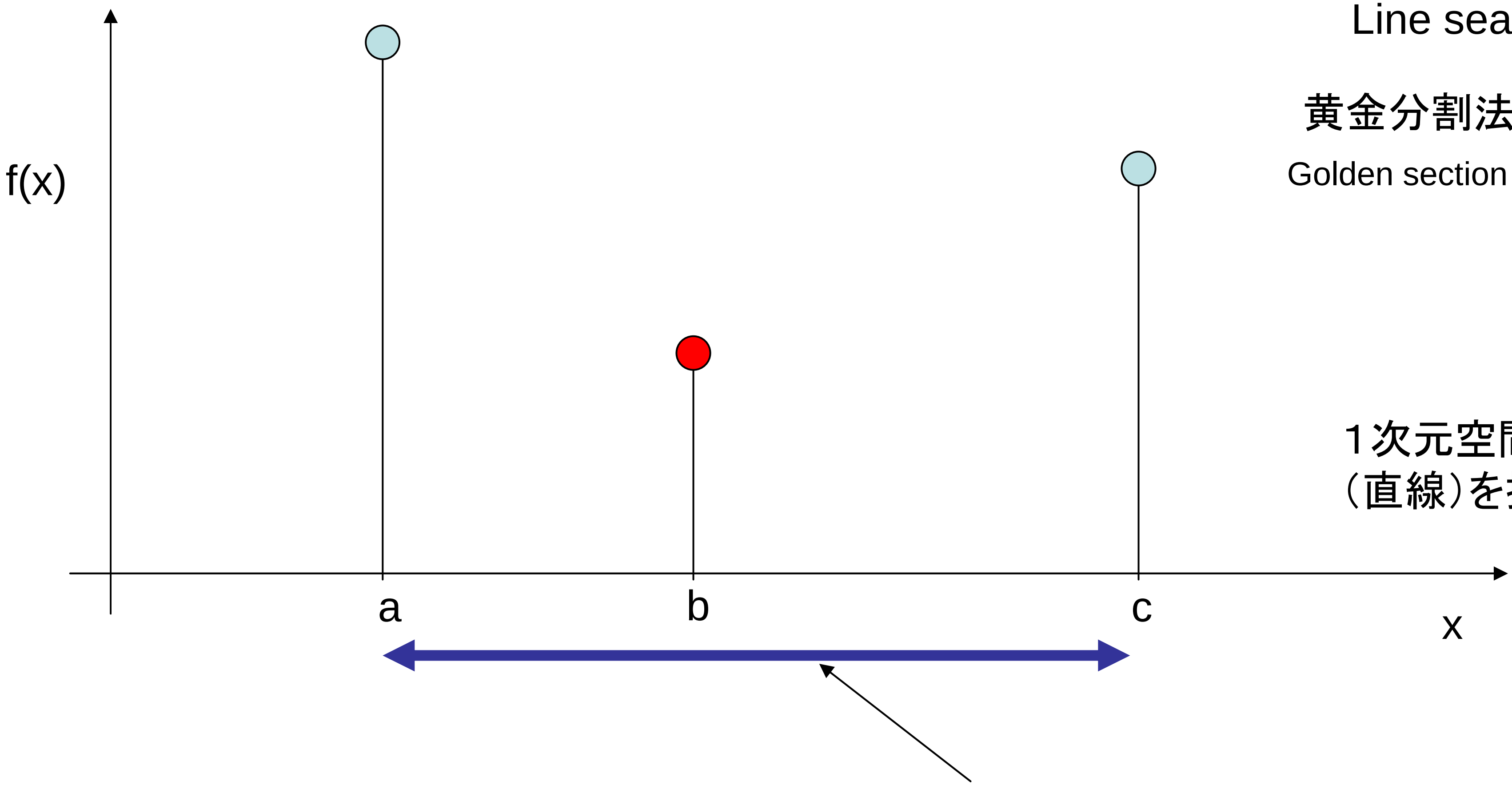
1） a < b < c の3点をとる　 f(b) が最小なら、区間 [a,c] に必ず極小がある。

2） [b, c] の区間で新しい点 x をとり、f(x) を計算する。

3） f(b) < f(x) なら、a < b < x に狭められた範囲に極小がある
　 f(x) < f(b) なら、b < x < c に狭められた範囲に極小がある

【制約のない関数最適化】1次元の最小化:ラインサーチ

Line search



黄金分割法

Golden section search

1次元空間上
(直線)を探索

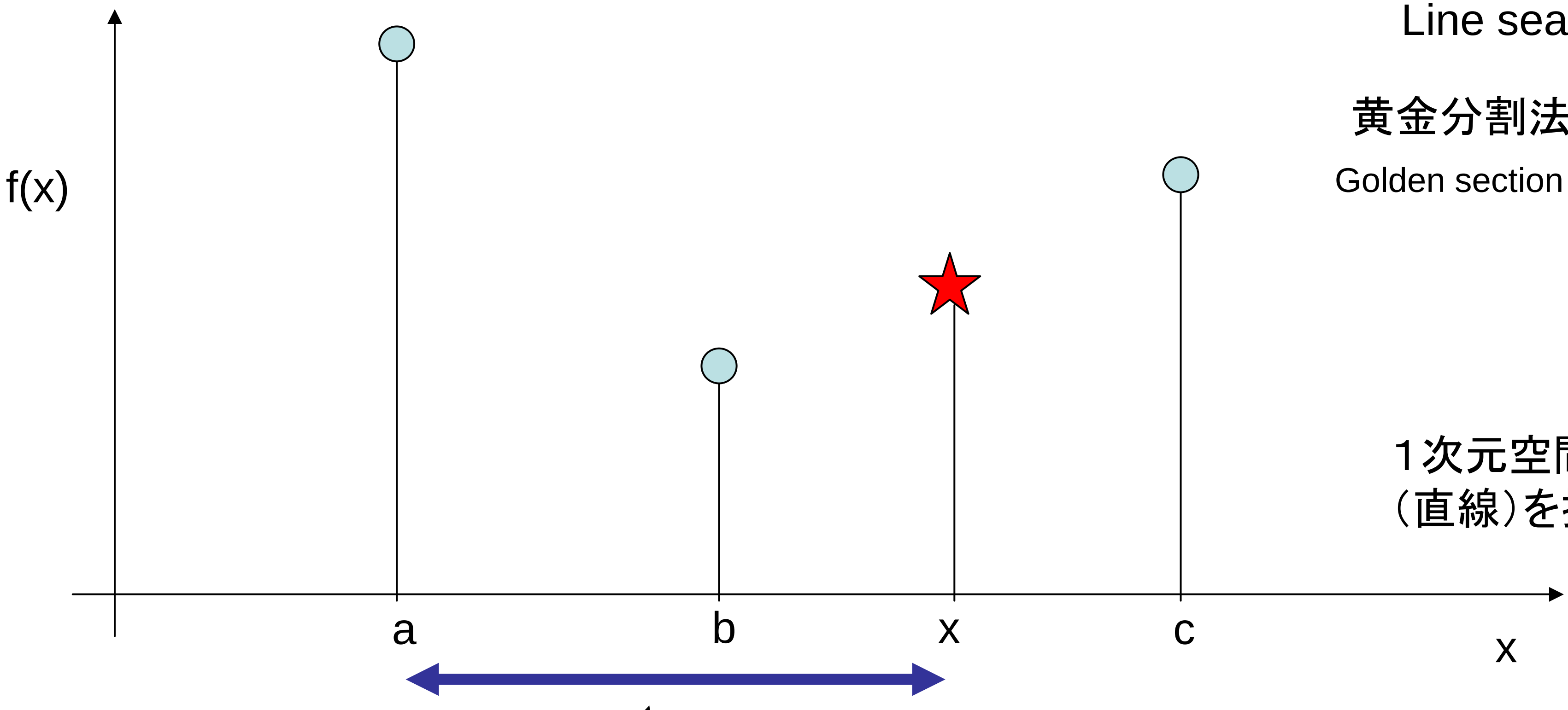
1） a < b < c の3点をとる　f(b) が最小なら、区間 [a,c] に必ず極小がある。

2） [b, c] の区間で新しい点 x をとり、f(x) を計算する。

3） f(b) < f(x) なら、a < b < x に狭められた範囲に極小がある
f(x) < f(b) なら、b < x < c に狭められた範囲に極小がある

【制約のない関数最適化】1次元の最小化：ラインサーチ

Line search



黄金分割法

Golden section search

1次元空間上
（直線）を探索

1） a < b < c の3点をとる　f(b) が最小なら、区間 [a,c] に必ず極小がある。

2） [b, c] の区間で新しい点 x をとり、f(x) を計算する。

3） f(b) < f(x) なら、a < b < x に狭められた範囲に極小がある→ a,b,x をa,b,c に置換
f(x) < f(b) なら、b < x < c に狭められた範囲に極小がある

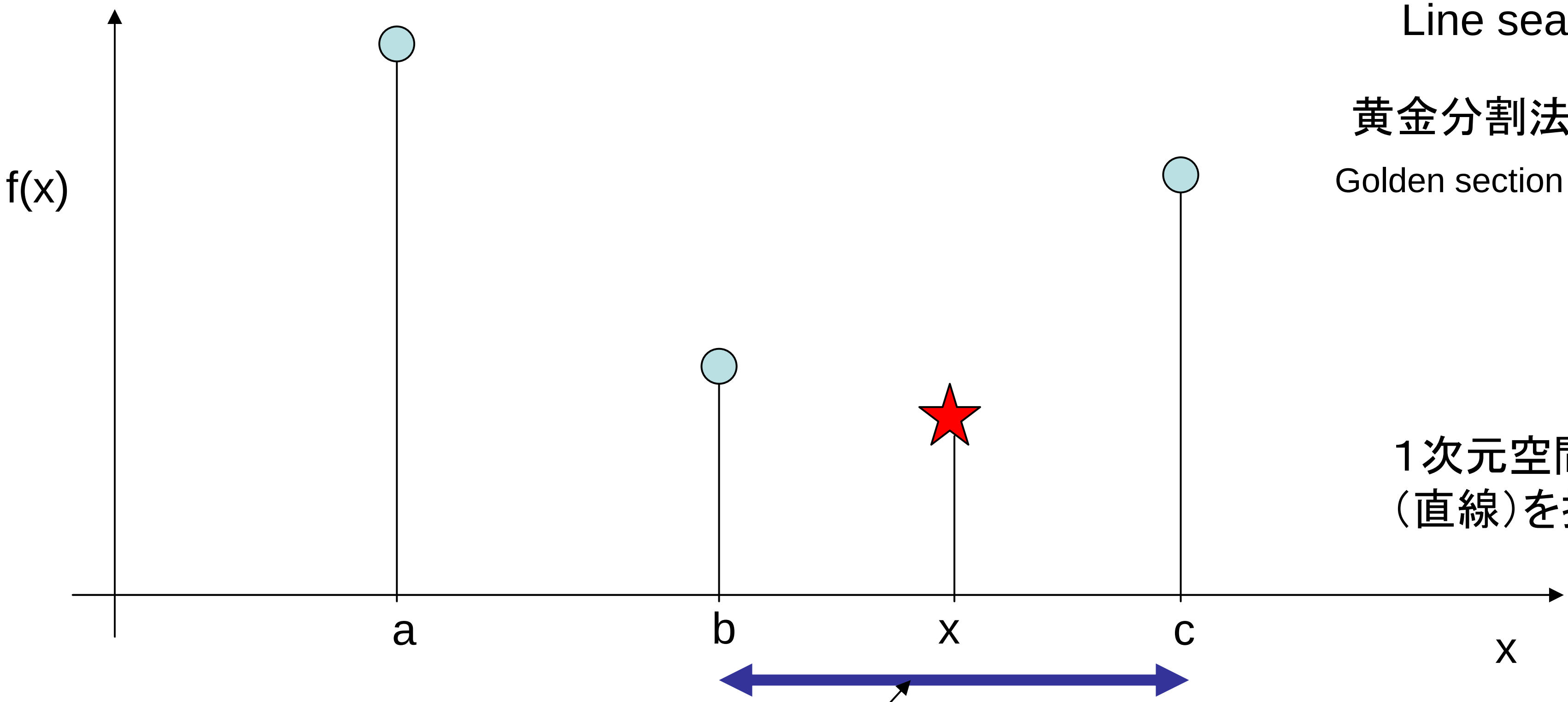
#【制約のない関数最適化】1次元の最小化:ラインサーチ

Line search

黄金分割法

Golden section search

1次元空間上
(直線)を探索



1） a < b < c の3点をとる　 f(b) が最小なら、区間 [a,c] に必ず極小がある。

2） [b, c] の区間で新しい点 x をとり、f(x) を計算する。

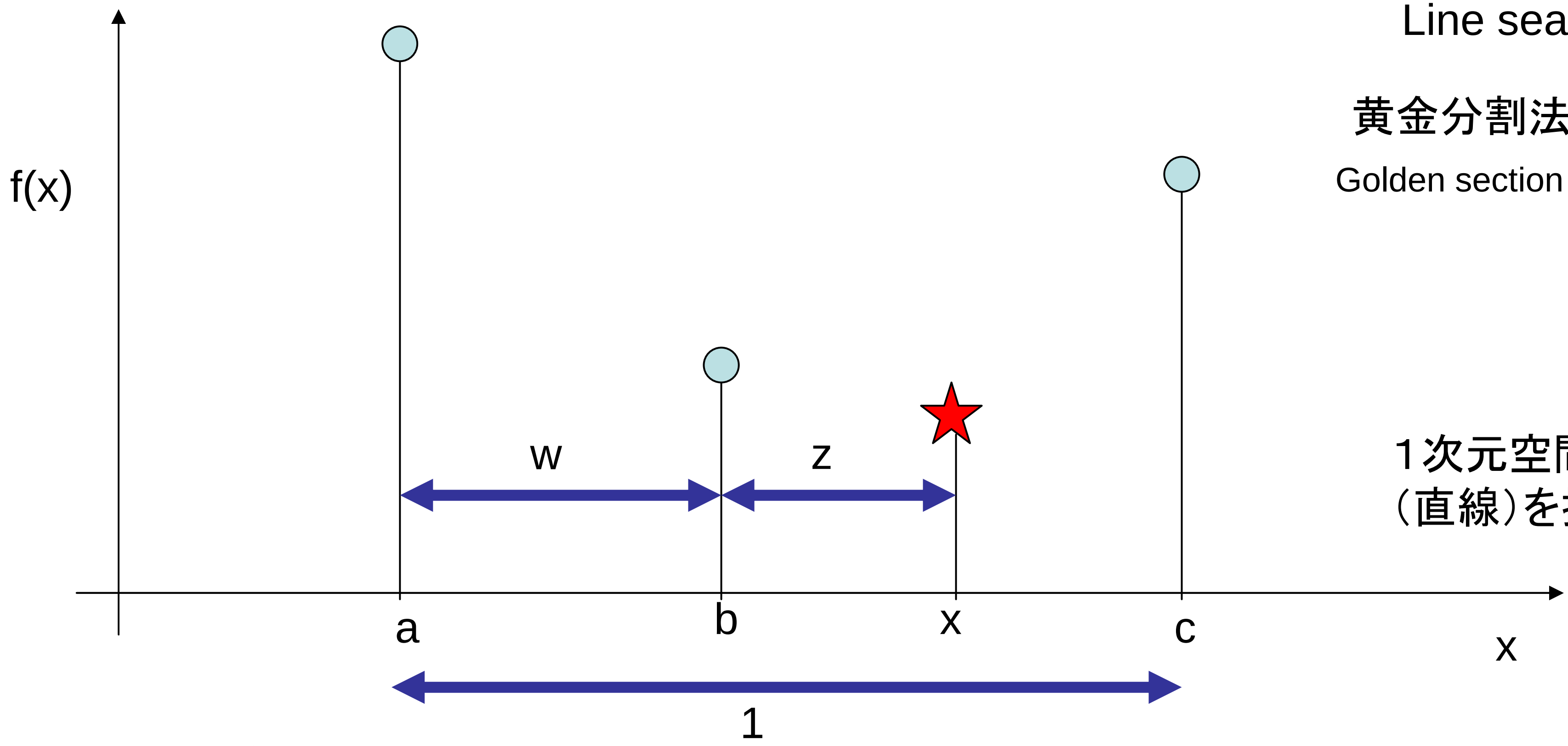
3） f(b) < f(x) なら、a < b < x に狭められた範囲に極小がある
f(x) < f(b) なら、b < x < c に狭められた範囲に極小がある→ b,x,c をa,b,c に置換

【制約のない関数最適化】1次元の最小化:ラインサーチ

Line search

黄金分割法

Golden section search



1次元空間上
(直線)を探索

b や x をどのように決めるか？

次の新しい区間長は w+z または 1-w 。これらを等しくすると z = 1 - 2w

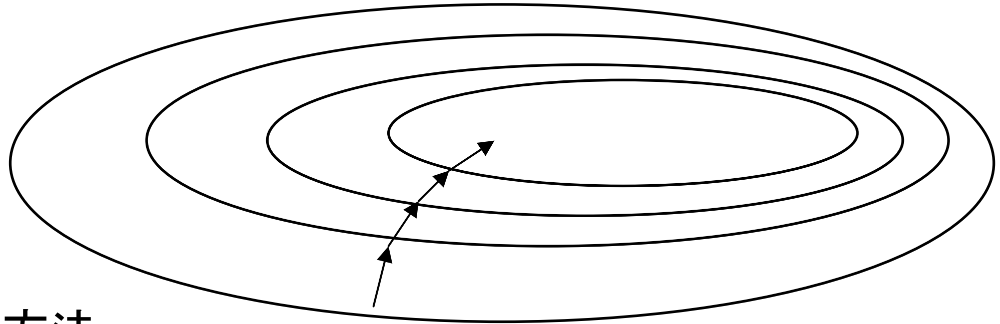
区間[b,c]に対するxの位置関係は、区間[a,c]に対するbの位置関係に等しい
1 : w = (1 – w) : z

これらの連立方程式を解くと、 $w^2 - 3w + 1 = 0$ $w = \frac{3-\sqrt{5}}{2}$ **黄金比** Golden section

【制約のない関数最適化】n次元関数の最小化

optimization in n-dimensional function

探索点での関数の値（と勾配の大きさ）に基づいて次の探索点を順次決定する手順により、あたかも山の頂上を目指すように、関数の極大値（あるいは極小値）をすみやかに発見する手法

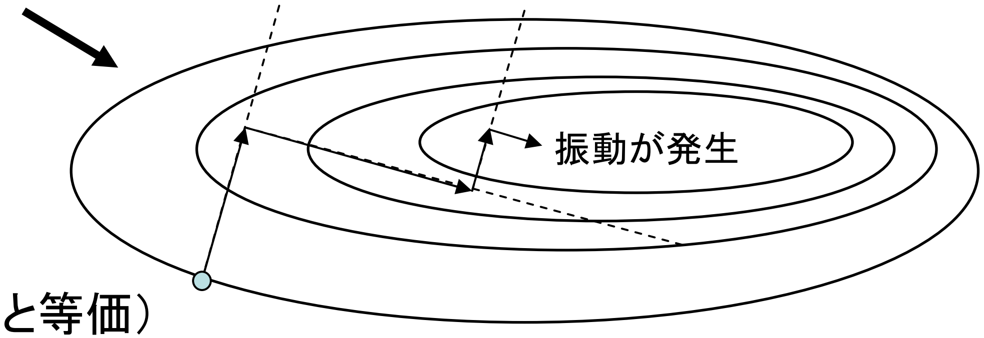


1）**勾配情報を用いる方法**：

・最大勾配法（最急降下法）
　勾配方向へ少しずつ解を改善

●ラインサーチと方向転換を繰り返す方法

・最適勾配法：勾配方向にラインサーチを行い、その極小点で再び勾配方向へラインサーチ
だいたいうまく行くが、２次関数において極めて効率の悪い場合がある



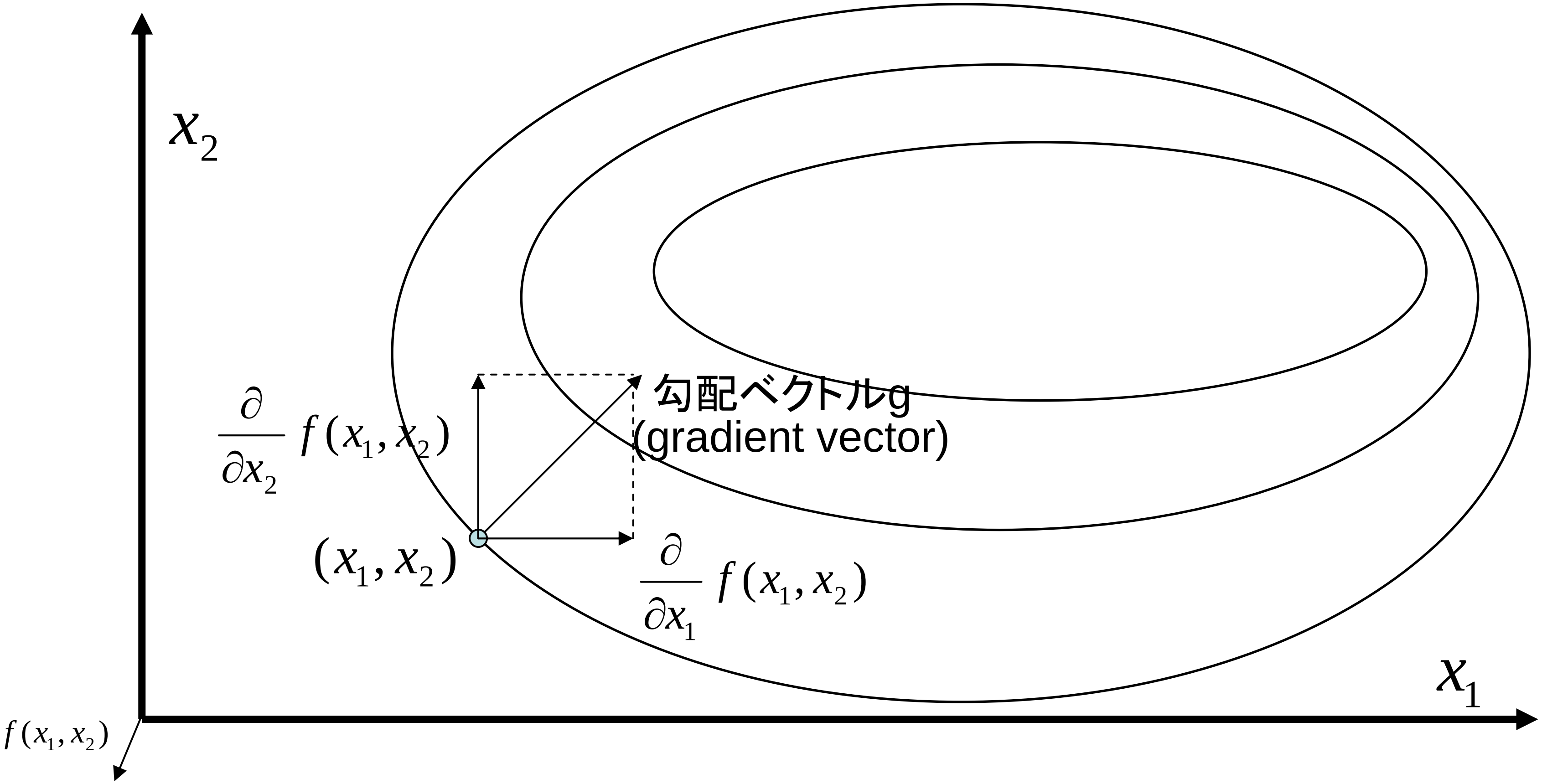
・共役勾配法：方向転換するとき、その点の勾配方向だけでなく、今まで下ってきた方向も考える
（関数の2階偏導関数を考慮に入れることと等価）
対象とする関数がn変数の２次形式である場合、
ラインサーチをn回繰返すことで最大値が求まることが保障される

【復習】 関数の勾配とは？
(gradient)

2変数関数 $f(x_1, x_2)$

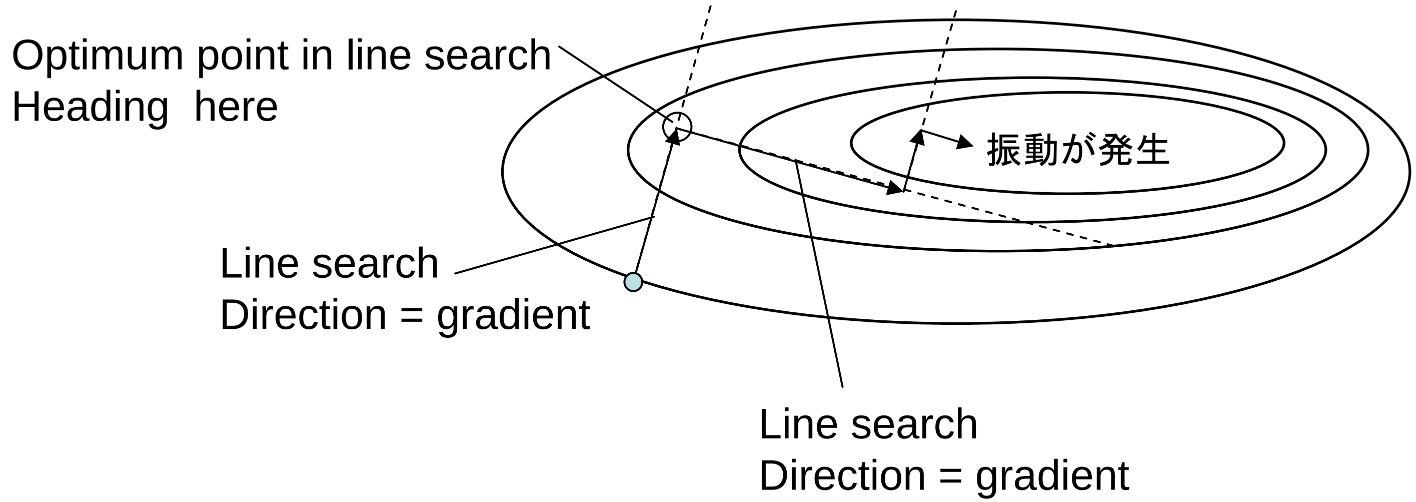
探索すべき
パラメータベクトル $\mathbf{x} = (x_1, x_2)$

勾配ベクトル
(gradient vector) $\mathbf{g} = \nabla f(x_1, x_2) = \left(\frac{\partial}{\partial x_1} f(x_1, x_2), \frac{\partial}{\partial x_2} f(x_1, x_2) \right)$



ラインサーチと方向転換を繰り返す方法（1）
最適勾配法（Optimal gradient method）

勾配方向にラインサーチを行い、極小点で再び勾配を求め、その方向へラインサーチする
Execute line search in direction of the gradient to its optimum point,
and turn to the new gradient direction, and execute line search again.



【問題点】
２次関数において、細かいステップで方向転換を繰り返しながら
同じような方向を何度もラインサーチを行い、谷底に着くまで多数のステップを要する
It needs many steps in quadratic functions repeating the line search
in similar direction.

新しい探索方向は、今までの全ての探索方向とも干渉しない（直交・共役である）
ことが望ましい A new direction should be conjugate against all past directions.

なぜ「2次形式」が重要なのか？
(quadratic)

ある特定の点Pを原点とし、この点の近傍座標をxとする。
すると、**どんな関数 f もテイラー級数で近似できる:**
(Taylor series)

$$f(\mathbf{x}) = f(\mathbf{P}) + \sum_i \frac{\partial f}{\partial x_i} x_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} x_i x_j + \cdots$$

$$= \underline{f(\mathbf{P}) + \nabla f|_{\mathbf{P}} \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}\cdot\mathbf{A}\cdot\mathbf{x}} + \cdots$$

2次形式
(quadratic)

この行列A はf(x)のHesse行列
(Hessian matrix)
各要素はPにおける2階編導関数

$$[\mathbf{A}]_{ij} \equiv \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathbf{P}}$$

【復習】「2次形式」と「共役」な方向とは？
(quadratic)　(conjugate)

$$f(x_1, x_2, \cdots, x_n) = \mathbf{X}^t \mathbf{A} \mathbf{X} + \mathbf{B}^t \mathbf{X} + C$$

2次形式 (quadratic)

ただし

探索すべきパラメータベクトル (parameter vector)

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Symmetrical matrix

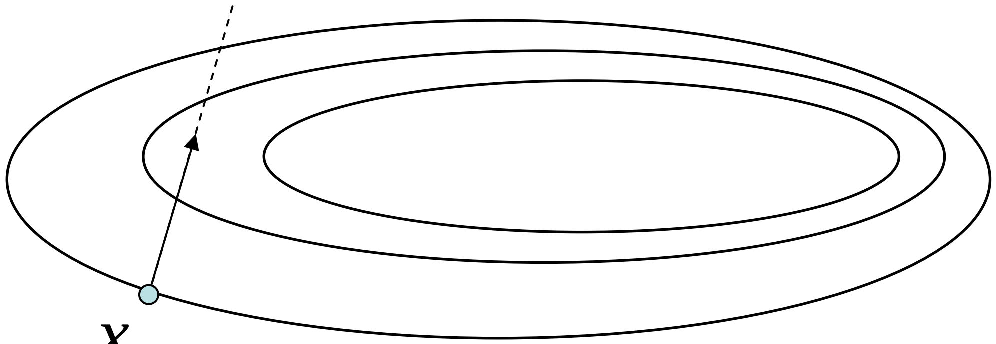
N×N 対称行列 $\mathbf{A}$ に対して、2つの方向を表すベクトル $\mathbf{p}$ と $\mathbf{q}$ が

$\mathbf{p}^t \mathbf{A} \mathbf{q} = 0$ を満たすとき、 $\mathbf{p}$ と $\mathbf{q}$ は $\mathbf{A}$ に関して互いに**共役**であるという
(conjugate)

参考 : $\mathbf{A}$ を単位行列とすると、上の式は直交条件となり、共役性は直交性の拡張概念

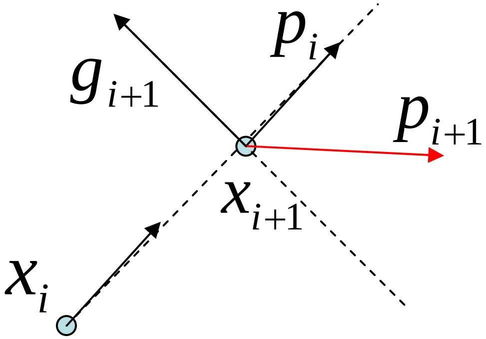
【制約のない関数最適化】共役勾配法 Conjugate gradient method

x_i	i 番目の点の座標ベクトル Coordinate of the i[th] search point
g_i	点 x_i における勾配ベクトル Gradient vector at x_i
p_i	点 x_i からラインサーチを行う方向ベクトル Direction of line search at x_i



まず最初は勾配の反対方向へラインサーチ

$$p_0 = -g_0$$



x_i の次の点 x_{i+1} を、ラインサーチで見つけた最小点とする

$$x_{i+1} = x_i + \alpha_i\, p_i$$ Optimum point in line search

次の点 x_{i+1} での新しい探索方向 p_{i+1} は、前の探索方向 p_i と共役であるようにする。
そのため、点 x_{i+1} での勾配と前の探索方向 p_i とを結合する:

$$p_{i+1} = -g_{i+1} + \beta_{i+1}\, p_i$$

ただし、重み係数 β_{i+1} は次の式のどちらかで与える:

$$\beta_{i+1} = \frac{g_{i+1}^T\, g_{i+1}}{g_i^T\, g_i} = \frac{\left\| g_{i+1} \right\|^2}{\left\| g_i \right\|^2}$$ （Fletcher-Reeves法）

（Polak-Ribiere法）

$$\beta_{i+1} = \frac{(g_{i+1} - g_i)^T\, g_{i+1}}{g_i^T\, g_i}$$

【共役勾配法が2次形式関数でn回のラインサーチで最適解を見つける理由】

次式を満たす n 個の n 次元ベクトル $\mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$ を考える:

$$\mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j = \begin{cases} const. & for \quad i = j \\ 0 & for \quad i \neq j \end{cases} \quad \Rightarrow \quad \mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$$ は共役ベクトル集合

次に、ベクトル $\mathbf{X}$ を $\mathbf{X} = \sum_{j=1}^{n} \alpha_j \mathbf{Z}_j$ と表すと、2次形式の評価式は

$$f(x_1, x_2, \cdots, x_n) = \mathbf{X}^t \mathbf{A} \mathbf{X} + \mathbf{B}^t \mathbf{X} + C$$

$$= \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j \right) + \left(\sum_{i=1}^{n} \alpha_i \mathbf{B}^t \mathbf{Z}_i \right) + C = \sum_{i=1}^{n} \alpha_i{}^2 \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_i + \alpha_i \mathbf{B}^t \mathbf{Z}_i + C$$

【共役勾配法が2次形式関数でn回のラインサーチで最適解を見つける理由】

次式を満たす n 個の n 次元ベクトル $\mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$ を考える:

$$\mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j = \begin{cases} const. & for \quad i = j \\ 0 & for \quad i \neq j \end{cases} \quad \Rightarrow \quad \mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n \text{ は共役ベクトル集合}$$

次に、ベクトル $\mathbf{X}$ を $\mathbf{X} = \sum_{j=1}^{n} \alpha_j \mathbf{Z}_j$ と表すと、2次形式の評価式は

$$f(x_1, x_2, \cdots, x_n) = \mathbf{X}^t \mathbf{A} \mathbf{X} + \mathbf{B}^t \mathbf{X} + C$$

$$= \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j \right) + \left(\sum_{i=1}^{n} \alpha_i \mathbf{B}^t \mathbf{Z}_i \right) + C = \sum_{i=1}^{n} \alpha_i{}^2 \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_i + \alpha_i \mathbf{B}^t \mathbf{Z}_i + C$$

ここで $g_i(\alpha_i) = \alpha_i{}^2 \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_i + \alpha_i \mathbf{B}^t \mathbf{Z}_i$ とおくと、上の式は以下のようになる:

各共役ベクトル毎にインデックス表示

【共役勾配法が2次形式関数でn回のラインサーチで最適解を見つける理由】

次式を満たす n 個の n 次元ベクトル $\mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$ を考える:

$$\mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j = \begin{cases} const. & for \quad i = j \\ 0 & for \quad i \neq j \end{cases} \Rightarrow \mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$$

は共役ベクトル集合

次に、ベクトル $\mathbf{X}$ を $\mathbf{X} = \sum_{j=1}^{n} \alpha_j \mathbf{Z}_j$ と表すと、2次形式の評価式は

$$f(x_1, x_2, \cdots, x_n) = \mathbf{X}^t \mathbf{A} \mathbf{X} + \mathbf{B}^t \mathbf{X} + C$$

$$= \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_j \right) + \left(\sum_{i=1}^{n} \alpha_i \mathbf{B}^t \mathbf{Z}_i \right) + C = \sum_{i=1}^{n} \alpha_i{}^2 \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_i + \alpha_i \mathbf{B}^t \mathbf{Z}_i + C$$

ここで $g_i(\alpha_i) = \alpha_i{}^2 \mathbf{Z}_i{}^t \mathbf{A} \mathbf{Z}_i + \alpha_i \mathbf{B}^t \mathbf{Z}_i$ とおくと、上の式は以下のようになる:

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} g_i(\alpha_i) + C \Rightarrow$$

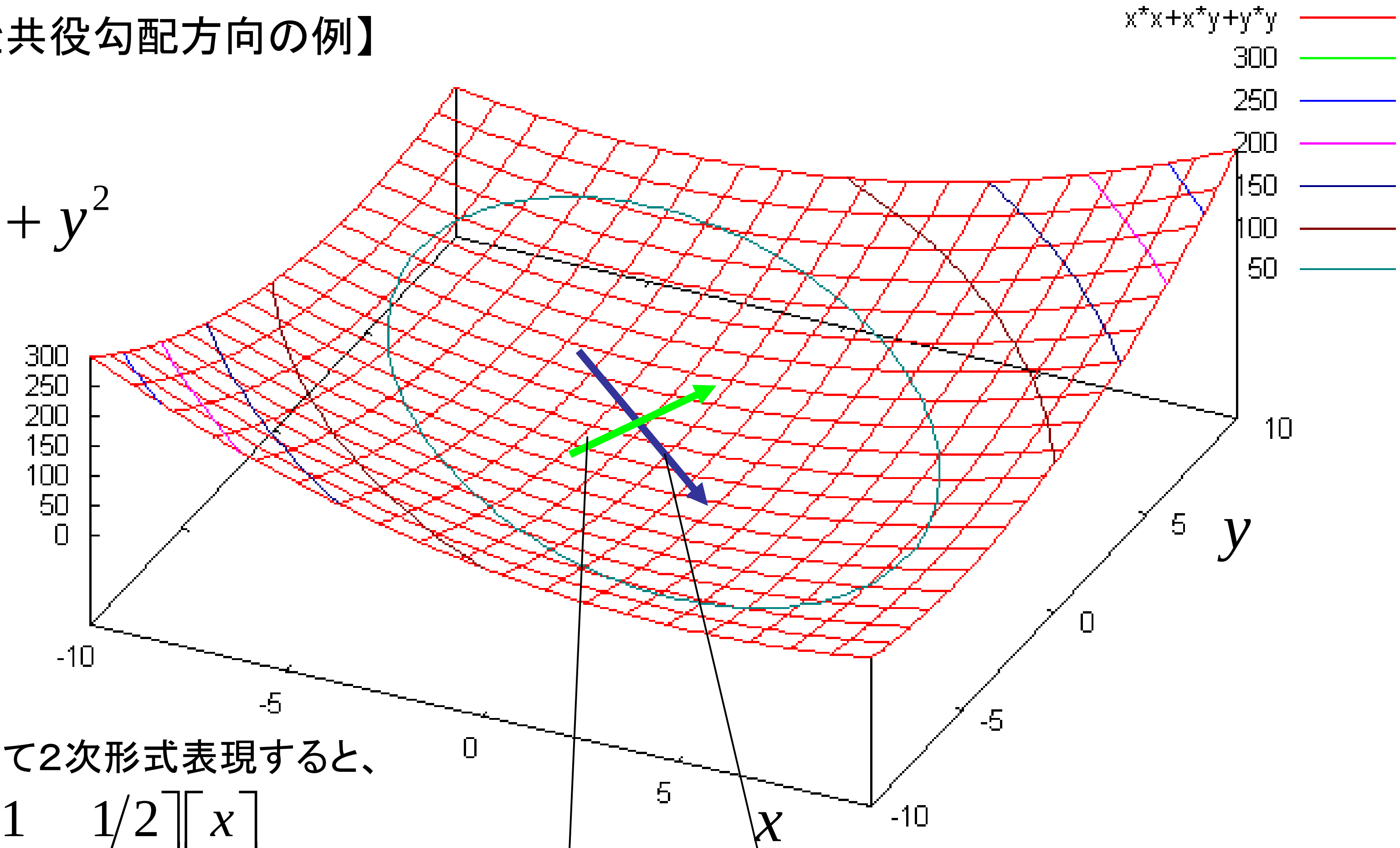
これより、関数 f はそれぞれ α_i に対する n 個の関数の代数和で表される

よって $g_1(\alpha_1), g_2(\alpha_2), \cdots g_n(\alpha_n)$ のそれぞれについて最小値を求めれば良い

探索開始点から順次 $\mathbf{Z}_1, \mathbf{Z}_2, \cdots \mathbf{Z}_n$ 方向の関数の断面について $\alpha_1, \alpha_2 \cdots \alpha_n$ の最小化を行えばn回のラインサーチで最適解を得る

【2次形式関数と共役勾配方向の例】

$$f = x^2 + xy + y^2$$



対称行列Aを用いて2次形式表現すると、

$$f = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

短軸方向

長軸方向

このとき

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0$$

となるので（1，1）と（1、－1）が対となる共役ベクトル