情報数理工学・コンピュータサイエンス実験第2
大規模連立1次方程式に対する共役勾配法

緒方秀教 (ogata@im.uec.ac.jp)

2016年11月2日（水）

実験スケジュール

- 11 月 2 日（水）
  - ガイダンス
  - 実験（計算機室，エリア 1）
- 10 月 14 日（水）~
  - 実験（計算機室，エリア 1）
- 実験内容
  - 課題 1~ 5 を解く→レポート
  - 証明問題（課題 1,2）…自宅学習
  - プログラミング（課題 3,4,5）…計算機室
  - 課題 6 以降は発展課題
- レポート締め切り…最終日から 3 週間後くらい（掲示に注意）

1. 序：本実験の目的

大規模連立 1 次方程式の数値解法である共役勾配法を理解し，プログラミングする．

$$Ax = \boldsymbol{b}, \quad \text{i.e.} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}.$$

（仮定）A は大規模対称疎行列である．

$$N \text{ 巨大}, \quad a_{ij} = a_{ji}, \quad \text{大部分の } a_{ij} = 0.$$

1. 序：問題の動機

科学技術計算（微分方程式の数値解法）
$\Longrightarrow$ 連立 1 次方程式の解の計算に帰着されることが多い.

1. 序：問題の動機

科学技術計算（微分方程式の数値解法）
$\Longrightarrow$ 連立 1 次方程式の解の計算に帰着されることが多い.

（例） 常微分方程式の 2 点境界値問題→差分法による近似解

$$\begin{cases} -\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2} = f(x) \quad a < x < b, \\ \qquad y(a) = y(b) = 0. \end{cases}$$

1. 序：問題の動機

科学技術計算（微分方程式の数値解法）
$\Longrightarrow$ 連立 1 次方程式の解の計算に帰着されることが多い.

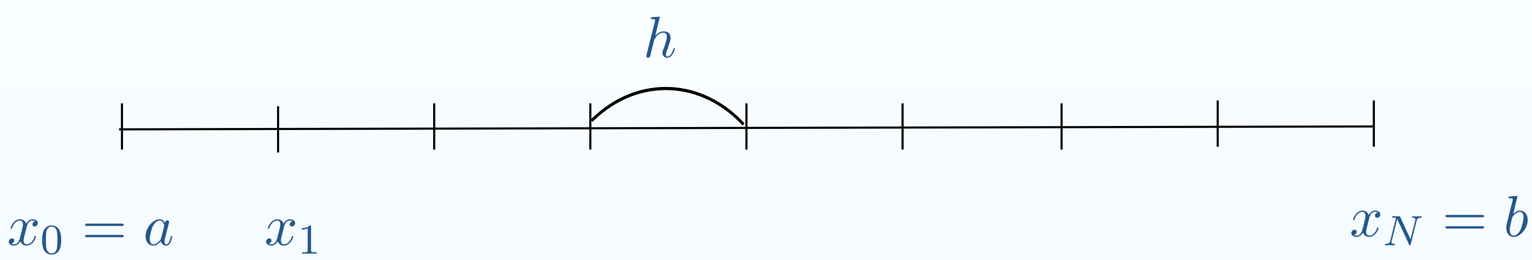
（例）常微分方程式の 2 点境界値問題→差分法による近似解

$$\begin{cases} -\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2} = f(x) \quad a < x < b, \\ y(a) = y(b) = 0. \end{cases}$$

差分法：微分→差分近似

$$\frac{\mathrm{d}y}{\mathrm{d}x} \simeq \frac{f(x+h) - f(x)}{h}, \quad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} \simeq \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

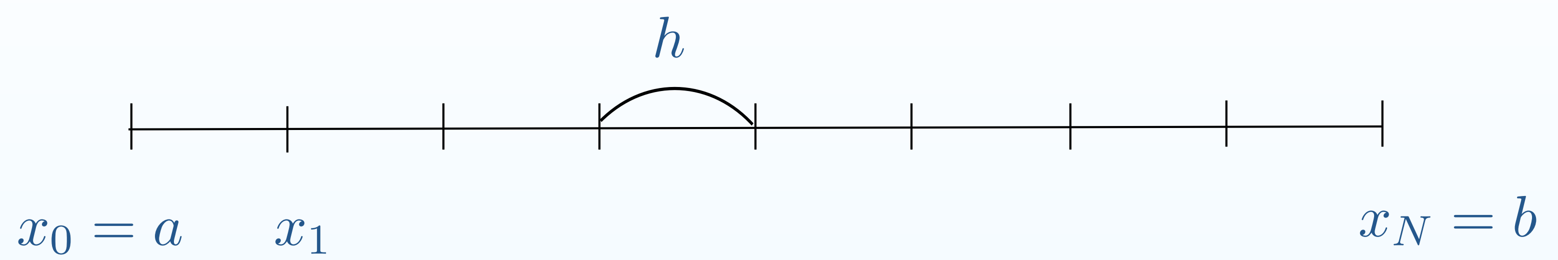
1. 序：問題の動機（微分方程式の差分法による近似解）



$$
\begin{bmatrix}
2 & -1 & & \\
-1 & 2 & -1 & \\
 & \ddots & \ddots & \ddots \\
 & & -1 & 2
\end{bmatrix}
\begin{bmatrix}
y_1 \\ y_2 \\ \vdots \\ y_{N-1}
\end{bmatrix}
=
\begin{bmatrix}
h^2 f(x_1) \\ h^2 f(x_2) \\ \vdots \\ h^2 f(x_{N-1})
\end{bmatrix},
$$

$$
\left.
\begin{aligned}
x_k &= a + kh \\
y_k &\simeq y(x_k)
\end{aligned}
\right\}
\quad k = 0, 1, \ldots, N;\ h = \frac{b-a}{N}.
$$

1. 序：問題の動機（微分方程式の差分法による近似解）



$$
\underbrace{\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{bmatrix}}_{\text{大規模疎行列}}
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \end{bmatrix}
=
\begin{bmatrix} h^2 f(x_1) \\ h^2 f(x_2) \\ \vdots \\ h^2 f(x_{N-1}) \end{bmatrix},
$$

$$
\left.\begin{aligned} x_k &= a + kh \\ y_k &\simeq y(x_k) \end{aligned}\right\} \quad k = 0, 1, \ldots, N;\ h = \frac{b-a}{N}.
$$

2. 理論

連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$

係数行列 A に対する仮定

1. **対称**：$A^T = A$
2. **正定値**：$(\boldsymbol{x}, A\boldsymbol{x}) \geq 0 \ (\ \forall \boldsymbol{x} \in \mathbb{R}^N\ ), \quad (\boldsymbol{x}, A\boldsymbol{x}) = 0 \ \Rightarrow \ \boldsymbol{x} = \boldsymbol{0}.$

$$* \quad (\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + \cdots + x_N y_N \quad \text{(ベクトルの内積)}$$

2. 理論

連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$

係数行列 A に対する仮定

1. **対称**：$A^T = A$
2. **正定値**：$(\boldsymbol{x}, A\boldsymbol{x}) \geq 0$ ($\forall \boldsymbol{x} \in \mathbb{R}^N$), $\quad (\boldsymbol{x}, A\boldsymbol{x}) = 0 \ \Rightarrow \ \boldsymbol{x} = \boldsymbol{0}$.

$*\quad (\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + \cdots + x_N y_N$ （ベクトルの内積）

連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解を求めることは，次の 2 次関数を最小にする $\boldsymbol{x} = \boldsymbol{x}^*$ を求めることと同値である．

$$f(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x}, A\boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) = \frac{1}{2}\sum_{i,j=1}^{N} a_{ij} x_i x_j - \sum_{i=1}^{N} b_i x_i.$$

2. 理論

連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$

係数行列 A に対する仮定

1. **対称**：$A^T = A$
2. **正定値**：$(\boldsymbol{x}, A\boldsymbol{x}) \geq 0 \ (\ \forall \boldsymbol{x} \in \mathbb{R}^N\ ), \quad (\boldsymbol{x}, A\boldsymbol{x}) = 0 \ \Rightarrow \ \boldsymbol{x} = \boldsymbol{0}.$

$$* \quad (\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + \cdots + x_N y_N \quad \text{（ベクトルの内積）}$$

連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解を求めることは，次の 2 次関数を最小にする $\boldsymbol{x} = \boldsymbol{x}^*$ を求めることと同値である．

$$f(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x}, A\boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) = \frac{1}{2}\sum_{i,j=1}^{N} a_{ij} x_i x_j - \sum_{i=1}^{N} b_i x_i.$$

2 次関数 $f(x)$ の最小化問題を考える.

2. 理論

$f(\boldsymbol{x})$ の最小化問題 $\Rightarrow$ **反復法**により解く.

$$\text{近似解列}\ \boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \ldots \to \boldsymbol{x}^*, \quad f(\boldsymbol{x}^*) = \min_{\boldsymbol{x}\in\mathbb{R}^N} f(\boldsymbol{x}).$$

近似解列 $\{\boldsymbol{x}_k\}$…次の算法により生成する.

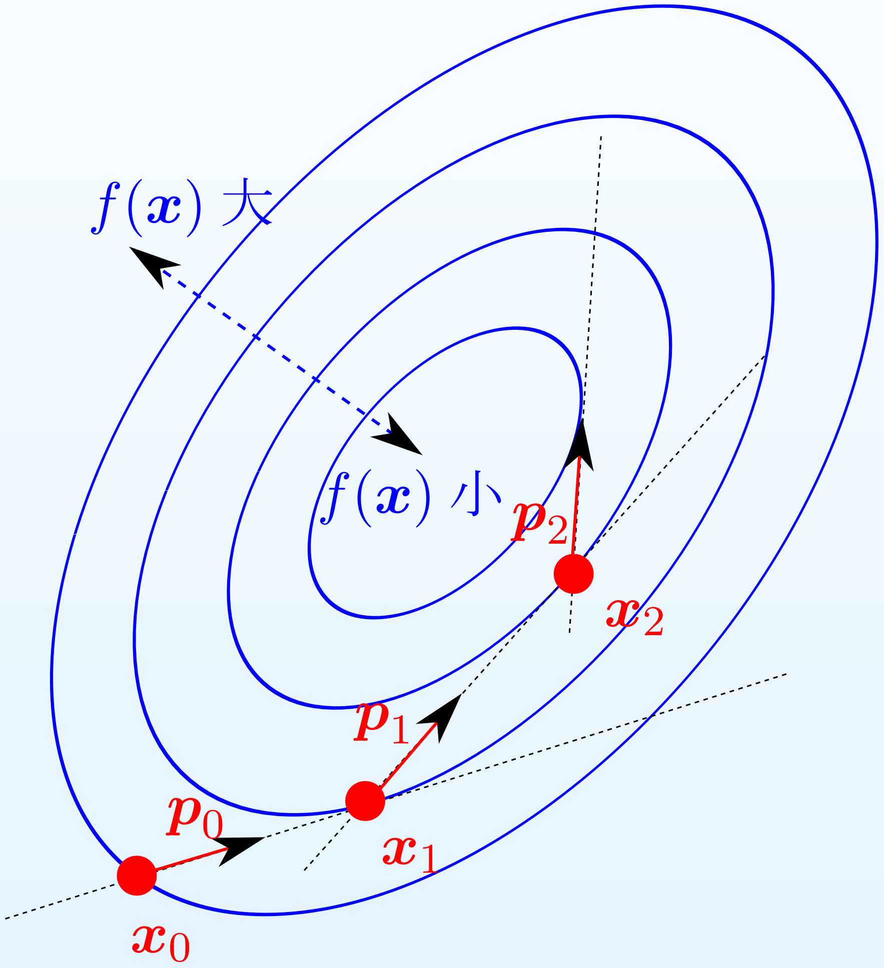
初期解 $\boldsymbol{x}_0$ を適当に定める；

$k = 0, 1, 2, \ldots$ に対して

- 探索方向ベクトル $\boldsymbol{p}_k$ を適当に定める；
- $f(\boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k) = \min_{\alpha\in\mathbb{R}} f(\boldsymbol{x} + \alpha \boldsymbol{p}_k)$ なる α_k を定める；
- $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k;$

2. 理論

$f(\boldsymbol{x})$ の等高線



2. 理論

$f(\boldsymbol{x})$ の最小化問題 $\Rightarrow$ **反復法**により解く.

$$\text{近似解列}\ \boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \ldots \to \boldsymbol{x}^*, \quad f(\boldsymbol{x}^*) = \min_{\boldsymbol{x}\in\mathbb{R}^N} f(\boldsymbol{x}).$$

近似解列 $\{\boldsymbol{x}_k\}$…次の算法により生成する.

初期解 $\boldsymbol{x}_0$ を適当に定める；

$k = 0, 1, 2, \ldots$ に対して

- 探索方向ベクトル $\boldsymbol{p}_k$ を適当に定める；
- $f(\boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k) = \min_{\alpha\in\mathbb{R}} f(\boldsymbol{x} + \alpha \boldsymbol{p}_k)$ なる α_k を定める；
- $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k$;

2 次関数 $f(\boldsymbol{x})$ に対しては，α_k は厳密に求まる（課題 1）.

$$\alpha_k = \frac{(\boldsymbol{p}_k, \boldsymbol{r}_k)}{(\boldsymbol{p}_k, A\boldsymbol{p}_k)}, \quad \boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k \quad \text{（残差ベクトル）}$$

2. 理論：最急降下法 (steepest descent method)

探索方向ベクトル $\boldsymbol{p}_k$ のとり方
…最急降下方向にとる ⇒ **最急降下法**

$$\boldsymbol{p}_k = -\nabla f(\boldsymbol{x}_k) = \boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k.$$

2. 理論：最急降下法 (steepest descent method)

探索方向ベクトル $\boldsymbol{p}_k$ のとり方
…最急降下方向にとる ⇒ **最急降下法**

$$\boldsymbol{p}_k = -\nabla f(\boldsymbol{x}_k) = \boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k.$$

問題点

同じ方向の探索方向ベクトル $\boldsymbol{p}_k$ が
何度も現れることがある.

$\Downarrow$

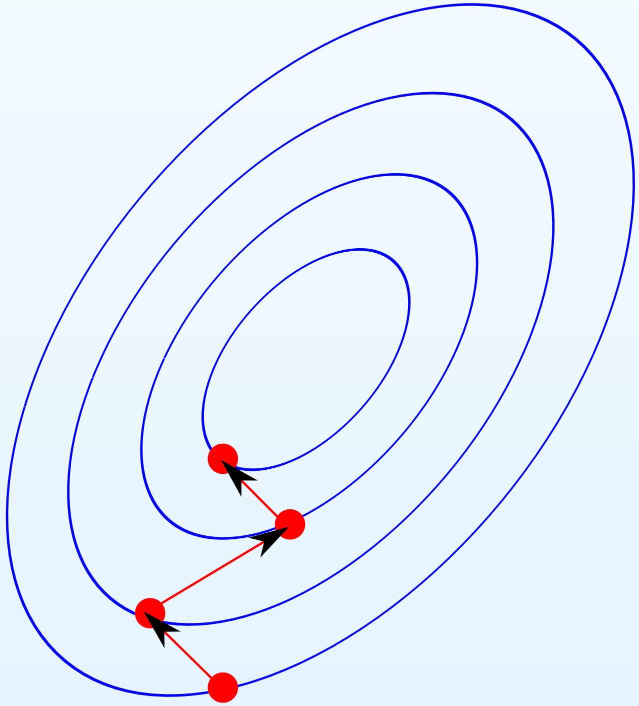
近似解を求める部分アフィン空間

$$\boldsymbol{x}_0 + \mathrm{span}\{\,\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_k\,\}$$

が広がらない.

$\Downarrow$

収束が遅くなる.



Stiefel の鳥かご

2. 理論：共役方向法 (conjugate direction method)

最急降下法の欠点を解消

$$A\text{-直交性} \quad (\boldsymbol{p}_j, A\boldsymbol{p}_k) = 0 \quad (\, j \neq k \,).$$

共役方向法$\cdots$探索方向ベクトル $\boldsymbol{p}_k$ が A-直交性を満たすようにとった反復法の総称.

2. 理論：共役方向法 (conjugate direction method)

最急降下法の欠点を解消

$$A\text{-直交性}\quad (\boldsymbol{p}_j, A\boldsymbol{p}_k) = 0 \quad (\, j \neq k\,).$$

共役方向法…探索方向ベクトル $\boldsymbol{p}_k$ が A-直交性を満たすようにとった反復法の総称.

探索方向ベクトルの系 $\{\, \boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_k \,\}$ は 1 次独立である（課題 2）.

（同じ方向の探索方向ベクトルを繰り返しとることはない）

2. 理論：共役方向法

探索方向ベクトル $\boldsymbol{p}_k$ の A-直交性により

直線探索（局所的最小化）の繰り返し

$$f(\boldsymbol{x}_{k+1}) = f(\boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k) = \min_{\alpha \in \mathbb{R}} f(\boldsymbol{x}_k + \alpha \boldsymbol{p}_k).$$

$$\Downarrow$$

大域的最適化の達成

$$\begin{aligned} f(\boldsymbol{x}_{k+1}) &= f(\boldsymbol{x}_0 + \alpha_0 \boldsymbol{p}_0 + \cdots + \alpha_k \boldsymbol{p}_k) \\ &= \min \left\{ \boldsymbol{x}_0 + c_0 \boldsymbol{p}_0 + \cdots + c_k \boldsymbol{p}_k \;\middle|\; c_0, c_1, \ldots, c_k \in \mathbb{R} \right\}. \end{aligned}$$

2. 理論：共役方向法

とくに $k = N - 1$ の場合,

$$\left\{ \boldsymbol{x}_0 + c_0 \boldsymbol{p}_0 + \cdots + c_{N-1} \boldsymbol{p}_{N-1} \mid c_0, \ldots, c_{N-1} \in \mathbb{R} \right\} = \mathbb{R}^N.$$

$$\Downarrow$$

$$f(\boldsymbol{x}_N) = \min_{\boldsymbol{x} \in \mathbb{R}^N} f(\boldsymbol{x}).$$

$\therefore \quad \boldsymbol{x}_N$ は $f(\boldsymbol{x})$ を最小にする，i.e., $A\boldsymbol{x} = \boldsymbol{b}$ の解である．

2. 理論：共役方向法

とくに $k = N-1$ の場合,

$$\left\{ \boldsymbol{x}_0 + c_0\boldsymbol{p}_0 + \cdots + c_{N-1}\boldsymbol{p}_{N-1} \mid c_0, \ldots, c_{N-1} \in \mathbb{R} \right\} = \mathbb{R}^N.$$

$$\Downarrow$$

$$f(\boldsymbol{x}_N) = \min_{\boldsymbol{x}\in\mathbb{R}^N} f(\boldsymbol{x}).$$

$\therefore \quad \boldsymbol{x}_N$ は $f(\boldsymbol{x})$ を最小にする，i.e., $A\boldsymbol{x} = \boldsymbol{b}$ の解である.

（理論上は）有限回の反復で連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の真の解 $\boldsymbol{x}^* = A^{-1}\boldsymbol{b}$ を得る.

2. 理論：共役方向法

$$
\begin{aligned}
f(\boldsymbol{x}_{k+1}) &= f(\boldsymbol{x}_0 + \alpha_0 \boldsymbol{p}_0 + \cdots + \alpha_k \boldsymbol{p}_k) \\
&= \min \{\, \boldsymbol{x}_0 + c_0 \boldsymbol{p}_0 + \cdots + c_k \boldsymbol{p}_k \mid c_0, \ldots, c_k \in \mathbb{R} \,\}
\end{aligned}
$$

の証明…下記のように「変数分離」できることから，証明できる．

$$
\begin{aligned}
& f(\boldsymbol{x}_0 + c_0 \boldsymbol{p}_0 + \cdots + c_k \boldsymbol{p}_k) \\
&= \frac{1}{2} \left(\boldsymbol{x}_0 + \sum_{j=0}^{k} c_k \boldsymbol{p}_k \right)^t A \left(\boldsymbol{x}_0 + \sum_{j=0}^{k} c_k \boldsymbol{p}_k \right) - \boldsymbol{b}^t \left(\boldsymbol{x}_0 + \sum_{j=0}^{k} c_k \boldsymbol{p}_k \right)
\end{aligned}
$$

($\{\boldsymbol{p}_k\}$ の A-直交性を用いて)

$$
= \sum_{j=0}^{k} \psi(c_j \boldsymbol{p}_j) + f(\boldsymbol{x}_0) \quad \cdots \quad c_0, \ldots, c_k \text{ について変数分離される}
$$

with $\quad \psi(\boldsymbol{x}) = \dfrac{1}{2} \boldsymbol{x}^t A \boldsymbol{x} - \boldsymbol{r}_0{}^t \boldsymbol{x}.$

2. 共役勾配法 (conjugate gradient method, CG 法)

A-直交性を満たす探索方向ベクトル $\boldsymbol{p}_k$ をどうやって生成するか？

2. 共役勾配法 (conjugate gradient method, CG 法)

A-直交性を満たす探索方向ベクトル $\boldsymbol{p}_k$ をどうやって生成するか？

最急降下方向ベクトル

$$-\nabla f(\boldsymbol{x}_k) = \boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k$$

を Gram-Schmidt 直交化してつくる．

$$\boldsymbol{p}_0 = \boldsymbol{r}_0,$$

$$\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} - \sum_{j=0}^{k} \beta_j^{(k)} \boldsymbol{p}_j, \quad \beta_j^{(k)} = \frac{(\boldsymbol{r}_{k+1}, A\boldsymbol{p}_j)}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)} \quad (\, k = 0, 1, \ldots\,).$$

こうして得られるのが**共役勾配法（CG 法）**である．

2. 理論：共役勾配法

実は次式が成立する．

$$\beta_j^{(k)} = \begin{cases} 0 & (\, j \leq k-1 \,) \\ -\|\boldsymbol{r}_{k+1}\|^2/\|\boldsymbol{r}_k\|^2 & (\, j = k \,) \end{cases}$$

$$\Downarrow$$

$$\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k, \quad \beta_k = \frac{\|\boldsymbol{r}_{k+1}\|^2}{\|\boldsymbol{r}_k\|^2}.$$

2. 理論：共役勾配法

実は次式が成立する.

$$\beta_j^{(k)} = \begin{cases} 0 & (\ j \le k-1\ ) \\ -\|\boldsymbol{r}_{k+1}\|^2/\|\boldsymbol{r}_k\|^2 & (\ j = k\ ) \end{cases}$$

$$\Downarrow$$

$$\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k, \quad \beta_k = \frac{\|\boldsymbol{r}_{k+1}\|^2}{\|\boldsymbol{r}_k\|^2}.$$

過去のステップの探索方向ベクトル $\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{k-1}$ を記憶しなくてよい（メモリの節約）.

2. 理論：共役勾配法

$$\beta_j^{(k)} = \begin{cases} 0 & (\ j \le k-1\ ) \\ -\|\boldsymbol{r}_{k+1}\|^2/\|\boldsymbol{r}_k\|^2 & (\ j = k\ ) \end{cases} \quad \text{の証明.}$$

Step 1/2　次の等式を証明する（プリント・命題 1）.

$$(\boldsymbol{r}_{k+1}, \boldsymbol{p}_j) = 0 \quad (\ j = 0, 1, \ldots, k\ ).$$

$j = k$ の場合,

$$\begin{aligned} 0 &= \left. \frac{\mathrm{d}}{\mathrm{d}\alpha} f(\boldsymbol{x} + \alpha \boldsymbol{p}_k) \right|_{\alpha=\alpha_k} = (\boldsymbol{p}_k, \nabla f(\boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k)) \\ &= (\boldsymbol{p}_k, A(\boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k)) - \boldsymbol{b} = -(\boldsymbol{p}_k, \boldsymbol{r}_{k+1}), \end{aligned}$$

$j = 0, \ldots, k-1$ の場合,

$$\begin{aligned} (\boldsymbol{r}_{k+1}, \boldsymbol{p}_j) &= (\boldsymbol{b} - A\boldsymbol{x}_{k+1}, \boldsymbol{p}_j) = (\boldsymbol{b} - A(\boldsymbol{x}_{j+1} + \alpha_{j+1}\boldsymbol{p}_{j+1} + \cdots + \alpha_{k+1}\boldsymbol{p}_{k+1}), \boldsymbol{p}_j) \\ &= \underbrace{(\boldsymbol{r}_{j+1}, \boldsymbol{p}_j)}_{0\ (\because\ j = k\ \text{の場合の証明})} \underbrace{-\alpha_{j+1}(\boldsymbol{p}_{j+1}, A\boldsymbol{p}_j) - \cdots - \alpha_{k+1}(\boldsymbol{p}_{k+1}, A\boldsymbol{p}_j)}_{0\ (\because\ \boldsymbol{p}_k\ \text{の}\ A\text{-直交性})} = 0. \end{aligned}$$

2. 理論：共役勾配法

Step 2/2　題意の等式の証明.

$$\beta_j^{(k)} = \frac{(\boldsymbol{r}_{k+1}, A\boldsymbol{p}_j)}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)} = \frac{1}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)}\left(\boldsymbol{r}_{k+1}, A\frac{\boldsymbol{x}_{j+1}-\boldsymbol{x}_j}{\alpha_j}\right) \quad (\because\ \boldsymbol{x}_{j+1} = \boldsymbol{x}_j + \alpha_j\boldsymbol{p}_j\ )$$

$$\left(\alpha_j = \frac{(\boldsymbol{r}_j, \boldsymbol{p}_j)}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)} = \frac{(\boldsymbol{r}_j, \boldsymbol{r}_j - \alpha_{j-1}\boldsymbol{p}_{j-1})}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)} \overset{(\because \text{Step 1/2})}{=} \frac{\|\boldsymbol{r}_j\|^2}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)} \text{ を用いて}\right)$$

$$= \frac{1}{(\boldsymbol{p}_j, A\boldsymbol{p}_j)}\frac{(\boldsymbol{p}_j, A\boldsymbol{p}_j)}{\|\boldsymbol{r}_j\|^2}(\boldsymbol{r}_{k+1}, \boldsymbol{r}_j - \boldsymbol{r}_{j+1}) = \frac{(\boldsymbol{r}_{k+1}, \boldsymbol{r}_j - \boldsymbol{r}_{j+1})}{\|\boldsymbol{r}_j\|^2}.$$

ここで,

$$\boldsymbol{r}_j = \boldsymbol{p}_j + \sum_{i=0}^{j-1}\beta_i^{(j-1)}\boldsymbol{p}_i \in \operatorname{span}\{\ \boldsymbol{p}_0, \ldots, \boldsymbol{p}_j\ \}, \quad (\boldsymbol{r}_{k+1}, \boldsymbol{p}_j) = 0 \quad (j = 0, \ldots, k)$$

を用いると，$j+1 \leq k$ $(j \leq k-1)$ の場合，$\beta_j^{(k)} = 0$.

$j+1 = k+1$ $(j = k)$ の場合，$\beta_j^{(k)} = -\dfrac{(\boldsymbol{r}_{k+1}, \boldsymbol{r}_{k+1})}{\|\boldsymbol{r}_k\|^2} = -\dfrac{\|\boldsymbol{r}_{k+1}\|^2}{\|\boldsymbol{r}_k\|^2}.$ ■

2. 理論：共役勾配法

共役勾配法のアルゴリズム

初期解 $\boldsymbol{x}_0$ を適当に定める； $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$; $\boldsymbol{p}_0 = \boldsymbol{r}_0$;

$k = 0, 1, 2, \ldots$ に対して

$$\alpha_k = \frac{\|\boldsymbol{r}_k\|^2}{(\boldsymbol{p}_k, A\boldsymbol{p}_k)};$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k;$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k A\boldsymbol{p}_k;$$

$\|\boldsymbol{r}_{k+1}\| \leq \epsilon \|\boldsymbol{b}\|$ ならば反復を終了する；

$$\beta_k = \frac{\|\boldsymbol{r}_{k+1}\|^2}{\|\boldsymbol{r}_k\|^2};$$

$$\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k;$$

3. ICCG 法：共役勾配法の歴史

1. 共役勾配法のはじまり
   - M. R. Hestenes & E. Stiefel:
     Methods of conjugate gradients for solving linear systems,
     Journal of Research of the National Bureau of Standards,
     49 (1952) 409–436.
   - （欠点）丸め誤差に弱い→長い間忘れ去られていた.
2. 1980 年代，再び日の目を見る.
   - 前処理との併用
   - Krylov 部分空間法としての再定式化
   - 非対称行列への拡張など（Bi-CG 法，etc.）

3. ICCG 法

前処理 (preconditioning)

もとの連立 1 次方程式の代わりに同値な別の連立 1 次方程式を解いて，共役勾配法の収束を速める．

3. ICCG 法

前処理 (preconditioning)

もとの連立 1 次方程式の代わりに同値な別の連立 1 次方程式を解いて，共役勾配法の収束を速める．

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\Downarrow$$

$$(C^{-1}AC^{-t})(C^{t}\boldsymbol{x}) = C^{-1}\boldsymbol{b}$$

3. ICCG 法

前処理 (preconditioning)

もとの連立 1 次方程式の代わりに同値な別の連立 1 次方程式を解いて，共役勾配法の収束を速める．

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\Downarrow$$

$$(C^{-1}AC^{-t})(C^{t}\boldsymbol{x}) = C^{-1}\boldsymbol{b}$$

$\widetilde{A} = C^{-1}AC^{-t}, \quad \widetilde{\boldsymbol{x}} = C^{t}\boldsymbol{x}, \quad \widetilde{\boldsymbol{b}} = C^{-1}\boldsymbol{b}$ とおくと，

$$\widetilde{A}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}} \quad \Leftarrow$$ これに共役勾配法を適用

3. ICCG 法

前処理 (preconditioning)

もとの連立 1 次方程式の代わりに同値な別の連立 1 次方程式を解いて，共役勾配法の収束を速める.

$$Ax = b$$

$$\Downarrow$$

$$(C^{-1}AC^{-t})(C^{t}x) = C^{-1}b$$

$\widetilde{A} = C^{-1}AC^{-t}, \quad \widetilde{x} = C^{t}x, \quad \widetilde{b} = C^{-1}b$ とおくと，

$$\widetilde{A}\widetilde{x} = \widetilde{b} \quad \Leftarrow \text{ これに共役勾配法を適用}$$

共役勾配法が速く収束するように行列 C を選ぶ.

3. ICCG 法：前処理

共役勾配法：係数行列が単位行列に近いほど速く収束する.

3. ICCG 法：前処理

共役勾配法：係数行列が単位行列に近いほど速く収束する.

定理　A の固有値を $\lambda_1, \lambda_2, \ldots, \lambda_N$ ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$) とすると,

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \leq \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} \cdot 4 \left(f(\boldsymbol{x}_0) - f(\boldsymbol{x}^*) \right)$$

$$\left(\boldsymbol{x}^* = A^{-1}\boldsymbol{b}, \quad \kappa = \frac{\lambda_1}{\lambda_N} \quad (\text{条件数}). \right)$$

□

3. ICCG 法：前処理

共役勾配法：係数行列が単位行列に近いほど速く収束する.

定理　A の固有値を $\lambda_1, \lambda_2, \ldots, \lambda_N$ ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$) とすると,

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \leq \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} \cdot 4 \left(f(\boldsymbol{x}_0) - f(\boldsymbol{x}^*) \right)$$

$$\left(\boldsymbol{x}^* = A^{-1}\boldsymbol{b}, \quad \kappa = \frac{\lambda_1}{\lambda_N} \quad \text{(条件数)}. \right)$$

□

$\kappa \simeq 1$ である（A が単位行列に近い）ほど速く収束する.

3. ICCG 法：前処理

共役勾配法：係数行列が単位行列に近いほど速く収束する.

定理　A の固有値を $\lambda_1, \lambda_2, \ldots, \lambda_N$ ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$) とすると,

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \leq \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k} \cdot 4\,(f(\boldsymbol{x}_0) - f(\boldsymbol{x}^*))$$

$$\left(\boldsymbol{x}^* = A^{-1}\boldsymbol{b}, \quad \kappa = \frac{\lambda_1}{\lambda_N} \quad \text{（条件数）.}\right)$$

$\square$

$\kappa \simeq 1$ である（A が単位行列に近い）ほど速く収束する.

$$\text{前処理} \quad C^{-1}AC^{-t}(C^t\boldsymbol{x}) = C^{-1}\boldsymbol{b}.$$

$$\Rightarrow \quad C^{-1}AC^{-t} \simeq I \quad \text{となるように行列 } C \text{ を選ぶ.}$$

3. ICCG 法：Cholesky 分解

$\boxed{\text{よく使われる前処理：不完全 Cholesky 分解}}$

対称行列 A に対する Cholesky 分解（LU 分解）

$$A = LDL^t = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ l_{N1} & l_{N2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_N \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{N1} \\ 0 & 1 & & l_{N2} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

3. ICCG 法：Cholesky 分解

よく使われる前処理：不完全 Cholesky 分解

対称行列 A に対する Cholesky 分解（LU 分解）

$$A = LDL^t = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ l_{N1} & l_{N2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_N \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{N1} \\ 0 & 1 & & l_{N2} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$\Downarrow$$

$$A = (LD^{1/2})(LD^{1/2})^t \quad \left(D^{1/2} = \begin{bmatrix} d_1^{1/2} & & \\ & \ddots & \\ & & d_N^{1/2} \end{bmatrix} \right),$$

$$\therefore \quad (LD^{1/2})^{-1} A (LD^{1/2})^{-t} = I.$$

3. ICCG 法：不完全 Cholesky 分解

不完全 Cholesky 分解 (Incomplete Cholesky decomposition)

$$a_{ij} = 0 \quad \Rightarrow \quad l_{ij} = 0$$

となるように，Cholesky 分解を省力化して行う.

3. ICCG 法：不完全 Cholesky 分解

不完全 Cholesky 分解 (Incomplete Cholesky decomposition)

$$a_{ij} = 0 \quad \Rightarrow \quad l_{ij} = 0$$

となるように，Cholesky 分解を省力化して行う．

$$A \simeq LDL^t = (LD^{1/2})(LD^{1/2})^t$$

$$\Downarrow$$

$$(LD^{1/2})^{-1} A (LD^{1/2})^{-t} \simeq I.$$

3. ICCG 法：不完全 Cholesky 分解

不完全 Cholesky 分解 (Incomplete Cholesky decomposition)

$$a_{ij} = 0 \quad \Rightarrow \quad l_{ij} = 0$$

となるように，Cholesky 分解を省力化して行う.

$$A \simeq LDL^t = (LD^{1/2})(LD^{1/2})^t$$

$$\Downarrow$$

$$(LD^{1/2})^{-1} A (LD^{1/2})^{-t} \simeq I.$$

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\Downarrow$$

$$\underbrace{(LD^{1/2})^{-1} A (LD^{1/2})^{-t}}_{\text{単位行列に近い}} (LD^{1/2})^t \boldsymbol{x} = (LD^{1/2})^{-1} \boldsymbol{b}.$$

3. ICCG 法

$$\widetilde{A} = (LD^{1/2})^{-1}(LD^{1/2})^{-t}, \quad \widetilde{\boldsymbol{x}} = (LD^{1/2})^{t}\boldsymbol{x}, \quad \widetilde{\boldsymbol{b}} = (LD^{1/2})^{-1}\boldsymbol{b}.$$

$$\Downarrow$$

$\widetilde{A}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}}$,　$\widetilde{A}$ は正方行列に近い行列.

3. ICCG 法

$$\widetilde{A} = (LD^{1/2})^{-1}(LD^{1/2})^{-t}, \quad \widetilde{\boldsymbol{x}} = (LD^{1/2})^{t}\boldsymbol{x}, \quad \widetilde{\boldsymbol{b}} = (LD^{1/2})^{-1}\boldsymbol{b}.$$

$$\Downarrow$$

$\widetilde{A}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}}$, $\widetilde{A}$ は正方行列に近い行列.

$A\boldsymbol{x} = \boldsymbol{b}$ の代わりに $\widetilde{A}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}}$ に共役勾配法（CG 法）を適用すれば,
より速く収束すると期待される.

$\cdots$ICCG 法 = IC 分解 + CG 法

3. ICCG 法

ICCG 法のアルゴリズム

初期解 $\boldsymbol{x}_0$ を適当に定める； $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0;\quad \boldsymbol{p}_0 = (LDL^t)^{-1}\boldsymbol{r}_0;$

$k = 0, 1, 2, \ldots$ に対して

$$\left[\begin{array}{l}
\alpha_k = \dfrac{(\boldsymbol{r}_k, (LDL^t)^{-1}\boldsymbol{r}_k)}{(\boldsymbol{p}_k, A\boldsymbol{p}_k)}; \\
\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k; \\
\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k A\boldsymbol{p}_k; \\
\|\boldsymbol{r}_{k+1}\| \leq \epsilon\|\boldsymbol{b}\| \text{ ならば反復を終了する；} \\
\beta_k = \dfrac{(\boldsymbol{r}_{k+1}, (LDL^t)^{-1}\boldsymbol{r}_{k+1})}{(\boldsymbol{r}_k, (LDL^t)^{-1}\boldsymbol{r}_k)}; \\
\boldsymbol{p}_{k+1} = (LDL^t)^{-1}\boldsymbol{r}_{k+1} + \beta \boldsymbol{p}_k;
\end{array}\right.$$

3. ICCG 法

ICCG 法のアルゴリズム

初期解 $\boldsymbol{x}_0$ を適当に定める； $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$; $\boldsymbol{p}_0 = (LDL^t)^{-1}\boldsymbol{r}_0$;

$k = 0, 1, 2, \ldots$ に対して

$$\left[\begin{array}{l}
\alpha_k = \dfrac{(\boldsymbol{r}_k, (LDL^t)^{-1}\boldsymbol{r}_k)}{(\boldsymbol{p}_k, A\boldsymbol{p}_k)}; \\
\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k; \\
\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k A\boldsymbol{p}_k; \\
\|\boldsymbol{r}_{k+1}\| \le \epsilon\|\boldsymbol{b}\| \text{ ならば反復を終了する；} \\
\beta_k = \dfrac{(\boldsymbol{r}_{k+1}, (LDL^t)^{-1}\boldsymbol{r}_{k+1})}{(\boldsymbol{r}_k, (LDL^t)^{-1}\boldsymbol{r}_k)}; \\
\boldsymbol{p}_{k+1} = (LDL^t)^{-1}\boldsymbol{r}_{k+1} + \beta \boldsymbol{p}_k;
\end{array}\right.$$

（注意）$\boldsymbol{y} = (LDL^t)^{-1}\boldsymbol{x}$ の計算は，次の $\boldsymbol{y}, \boldsymbol{z}$ に対する連立方程式を解いて行う（逆行列の計算をしない）．

$$\begin{cases} L\boldsymbol{z} = \boldsymbol{x} \\ L^t\boldsymbol{y} = D^{-1}\boldsymbol{z}. \end{cases}$$

まとめ

- 連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ に対する共役勾配法
  - 2 次関数 $f(\boldsymbol{x}) = (1/2)(\boldsymbol{x}, A\boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x})$ の最小化問題に反復法を適用.
  - 探索方向ベクトル⋯A-直交性.
  - A-直交性を満たす探索方向ベクトルを，最急降下方向ベクトルの Gram-Schmidt 直交化によりつくる.
- ICCG 法⋯不完全 Cholesky 分解による前処理.
- 参考文献
  - 杉原正顯，室田一雄：【岩波数学叢書】線形計算の数理, 岩波書店，2009 年（6000 円＋税）.