

連立一次方程式の数値解法 2 : 共役勾配法.

$$Ax = b$$

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

A : 大規模疎行列 (n : 巨大, A の非零成分の割合が小さい) ... 実用上の理由.

LU分解による解法は不適 (LU分解の計算量 $O(n^3)$)
別の解法が必要

(仮定) A : 正定値対称行列

$$(x, Ax) \geq 0 \quad (\forall x \in \mathbb{R}^n) \text{ かつ } (x, Ax) = 0 \Rightarrow x = 0, \quad A^T = A$$

$$\text{つまり, } (x, y) = \sum_{i=1}^n a_i x_i y_i \quad \left(x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) \text{ -- 内積}$$

連立一次方程式 $Ax = b$ の解を求めると、次の二次関数 $f(x)$ の最小化問題と同等である:

$$f(x) = \frac{1}{2}(x, Ax) - (b, x).$$

すなわち.

$$Ax^* = b \iff f(x^*) = \min_{x \in \mathbb{R}^n} f(x).$$

$$\left(\because f(x) = \frac{1}{2}(x - x^*, A(x - x^*)) + \text{const.} \quad \square \right)$$

< 反復解法 >

$x_0 \in \mathbb{R}^n$ を適当にと;

$k = 0, 1, 2, \dots$ に対して

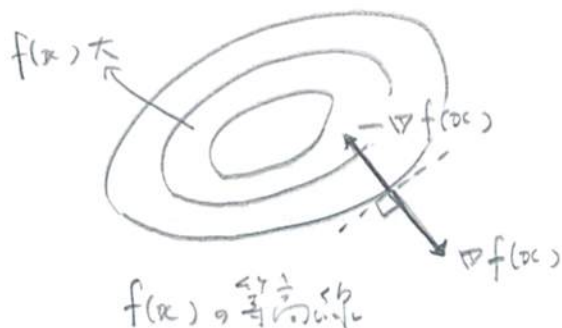
$$\left[\begin{array}{l} \text{探索方向 } p_k \text{ を } k \text{ に } \|p_k\| \text{ を適切に定めて;} \\ f(x_k) = \min_{\alpha \in \mathbb{R}} f(x_k + \alpha p_k) \text{ なる } \alpha_k \text{ を見つける;} \\ x_{k+1} = x_k + \alpha_k p_k; \end{array} \right.$$

$\rightarrow x_k \rightarrow x^*$ ($Ax = b$ の解) と期待される.

< 最急降下法 > 探索方向ベクトル $\|p_k$ をどう選ぶ?

素朴な思いつき法 ... $f(x)$ の最急降下方向は $\|p_k$ を選ぶ.

$$\|p_k = -\nabla f(x_k), \text{ where } \nabla f(x) = \begin{bmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_m \end{bmatrix}$$



u の場合, $-\nabla f(x) = r = b - Ax \dots$ 残差ベクトル

算法 ~ 最急降下法 ~

x_0 を適当にとる;

$$r_0 = \|p_0 = b - Ax_0;$$

$k = 0, 1, 2, \dots$ に対して

$$\left[\begin{array}{l} \alpha_k \text{ を } f(x_k + \alpha \|p_k) = \min_{\alpha \in \mathbb{R}} f(x_k + \alpha \|p_k) \text{ なる } \alpha_k \text{ を選ぶ;} \\ x_{k+1} = x_k + \alpha_k \|p_k; \\ r_{k+1} = r_k - \alpha_k A \|p_k; \\ \|p_{k+1} = r_{k+1}; \end{array} \right.$$

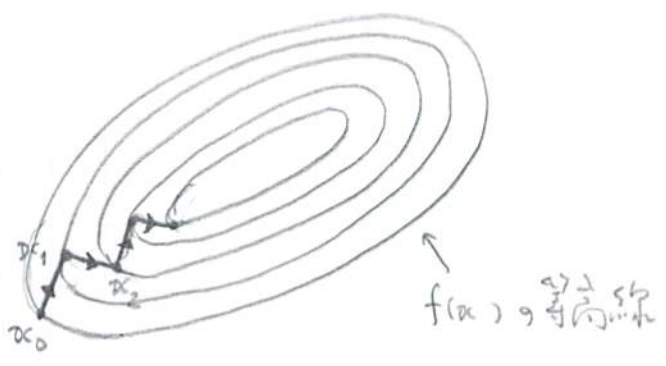
u の場合, $f(x)$ は二次関数である, α_k は厳密にとる:

$$\alpha_k = \frac{(r_k, \|p_k)}{(\|p_k, A \|p_k)}$$

$$\begin{aligned} \left(\begin{array}{l} \because f(x_k + \alpha \|p_k) = \frac{1}{2} (x_k + \alpha \|p_k, A(x_k + \alpha \|p_k)) - (b, x_k + \alpha \|p_k) \\ = \frac{1}{2} (\|p_k, A \|p_k) \alpha^2 - (b - Ax_k, \|p_k) \alpha + \text{const.} \\ = \frac{1}{2} (\|p_k, A \|p_k) \alpha^2 - (r_k, \|p_k) \alpha + \text{const.} \\ = \frac{1}{2} (\|p_k, A \|p_k) \left[\alpha - \frac{(r_k, \|p_k)}{(\|p_k, A \|p_k)} \right]^2 + \text{const.} \end{array} \right. \quad \square \end{aligned}$$

最急降下法の問題点

... 同じ方向の $\|p_k$ の何層も現れる可能性もある。
 したがって、解を探索する空間が狭くなる。
 p_k の解 x^* に近づきにくいことがある。



最急降下法の反復解 x_k の振舞い

< 共役方向法 > 最急降下法の問題点を解決する。

探索方向 n つに達した $\{p_0, p_1, \dots, p_{k-1}\}$ が一次独立な方向になる。
 したがって、解空間 S_k が正確に狭くなる。 $x_k \rightarrow x^*$ となる。

$$\boxed{(p_k, A p_r) = 0 \quad (k \neq r)} \quad \underline{A\text{-直交性}} \quad \text{---} (*)$$

このように p_k となる。

共役方向法 ... (*) を満たすような p_k を用いた反復解法の総称。

(*) が成立 $\rightarrow \{p_0, p_1, \dots, p_{k-1}\}$ は一次独立である。
 (各自直交している。)

このとき、次式が成立する (変数分離性) :

$$\left. \begin{aligned} f\left(x_0 + \sum_{j=0}^{k-1} c_j p_j\right) &= \sum_{j=0}^{k-1} \psi(c_j, p_j) + f(x_0), \\ \text{where } \psi(y_1) &= \frac{1}{2} (y_1, y_1) - (r_0, y_1). \end{aligned} \right\} \textcircled{*}$$

$$\begin{aligned}
(\because f(x_0 + \sum_{j=0}^{k-1} c_j p_j) &= \frac{1}{2} (x_0 + \sum_{j=0}^{k-1} c_j p_j, A(x_0 + \sum_{j=0}^{k-1} c_j p_j)) - (b, x_0 + \sum_{j=0}^{k-1} c_j p_j) \\
&= [(x) \quad z(A) \quad u] \\
&= \frac{1}{2} \sum_{j=0}^{k-1} c_j^2 (p_j, A p_j) - \sum_{j=0}^{k-1} c_j (b - A x_0, p_j) + f(x_0) \\
&= \frac{1}{2} \sum_{j=0}^{k-1} c_j (p_j, A p_j) - \sum_{j=0}^{k-1} c_j (r_0, p_j) + f(x_0) \\
&= \sum_{j=0}^{k-1} \psi(c_j | p_j) + f(x_0). \quad \square)
\end{aligned}$$

⊙ 27 .

$$\begin{aligned}
f(x_k) &= \min_{c \in \mathbb{R}} f(x_{k-1} + c | p_{k-1}) \\
&= \min \left\{ f(x_0 + \sum_{j=0}^{k-1} c_j | p_j) \mid c_0, c_1, \dots, c_{k-1} \in \mathbb{R} \right\}.
\end{aligned}$$

直線探索の系集を通る空間 $S_k = \{ x_0 + \sum_{j=0}^{k-1} c_j | p_j \}$ 全体での最小値を v_k とする。

と $k = n$ (空間の次元) とすれば、 $S_n = \mathbb{R}^n$ となる。

$$\begin{aligned}
f(x_n) &= \min \left\{ f(x_0 + \sum_{j=0}^{n-1} c_j | p_j) \mid c_0, c_1, \dots, c_{n-1} \in \mathbb{R} \right\} \\
&= \min_{x \in \mathbb{R}^n} f(x)
\end{aligned}$$

$\therefore x_n = x^*$

(理論上は) 有限回の反復で真の解 x^* に到達する。

< 共役勾配法 (conjugate gradient method, CG法) >

2nd. $(p_k, A p_k) = 0$ ($k \neq l$) 2nd $\{p_k\}$ 2nd 2nd 2nd 2nd 2nd?

→ 最急降下法, 探索方向の連続 $\{r_k\}$ "Gram-Schmidt 直交化" を施すこと.

($(x, y)_A \equiv (x, Ay)$ を内積と定義 (2nd の Gram-Schmidt 直交化)

$p_0 = r_0,$

$k = 0, 1, \dots, (n-1)$ まで

$$p_{k+1} = r_{k+1} - \sum_{j=0}^k \beta_j^{(k)} p_j \quad \text{where } \beta_j^{(k)} = \frac{(r_{k+1}, A p_j)}{(p_j, A p_j)}$$

$$(j = 0, 1, \dots, k).$$

↓
 実数 $\beta_0^{(k)} = \beta_1^{(k)} = \dots = \beta_{k-1}^{(k)} = 0$ とする.

→ 反復の間に, 直交化の p_0, p_1, \dots, p_{k-1} 2nd 2nd 2nd 2nd 2nd.

∴ (1) 3rd $(r_{k+1}, p_j) = 0$ ($j = 0, 1, \dots, k$) 2nd 成り立つ.

$$\because 0 = \frac{d}{d\alpha} f(x_k + \alpha p_k) \Big|_{\alpha = \alpha_k} = (p_k, \nabla f(x_k + \alpha_k p_k))$$

$$= (p_k, \nabla f(x_{k+1})) = -(p_k, r_{k+1}).$$

より, $(r_{k+1}, p_k) = 0$ 2nd 成り立つ.

2nd, $j \leq k-1$ まで, $\{p_k\}$ の A-直交性より

$$(r_{k+1}, p_j) = (r_{j+1} - \sum_{i=j+1}^k \alpha_i A p_i, p_j)$$

$$= \underbrace{(r_{j+1}, p_j)}_0 - \sum_{i=j+1}^k \alpha_i \underbrace{(p_i, A p_j)}_0 = 0.$$

(2) $\alpha_k \neq 0$ ($k = 0, 1, \dots, n-1$) と仮定して.

$$\because \alpha_k = \frac{(r_k, p_k)}{(p_k, A p_k)}, \quad (r_k, p_k) = (r_k, r_k - \sum_{j=0}^{k-1} \beta_j^{(k-1)} p_j) = \|r_k\|^2.$$

$r_k = 0$ 7d3 $x_k = x^*$ 2nd あり, 反復終了した. ☑

$$\begin{aligned} \beta_j^{(k)} \text{ の } 1 \text{ 次項} &= (r_{k+1}, A p_j) \\ &= \left(r_{k+1}, A \frac{x_j - x_{j-1}}{\alpha_{j-1}} \right) = \frac{1}{\alpha_{j-1}} (r_{k+1}, A x_j - A x_{j-1}) \\ &= \frac{1}{\alpha_{j-1}} (r_{k+1}, (b - A x_{j-1}) - (b - A x_j)) = \frac{1}{\alpha_{j-1}} \left\{ (r_{k+1}, r_{j-1}) - (r_{k+1}, r_j) \right\}. \end{aligned}$$

$j \leq k-1$ ならば

$$\begin{aligned} (r_{k+1}, r_j) &= \left(r_{k+1}, p_j + \sum_{i=0}^{j-1} \beta_i^{(j-1)} p_i \right) \\ &= \underbrace{(r_{k+1}, p_j)}_0 + \sum_{i=0}^{j-1} \beta_i^{(j-1)} \underbrace{(r_{k+1}, p_i)}_0 = 0, \end{aligned}$$

(i) $\forall j \leq k-1$ $(r_{k+1}, r_{j-1}) = 0$ ($j \leq k-1$)

ゆえに $(r_{k+1}, A p_j) = 0$ ($j \leq k-1$) $\therefore \beta_j^{(k)} = 0$ ($j \leq k-1$). □

$$\beta_k^{(k)} = \frac{(r_{k+1}, A p_k)}{(p_k, A p_k)} = \frac{1}{\alpha_k} \frac{(r_{k+1}, r_k - r_{k+1})}{(p_k, A p_k)} = \frac{(r_{k+1}, r_k) - \|r_{k+1}\|^2}{(r_k, p_k)}$$

$$(r_{k+1}, r_k) = (r_{k+1}, p_k + \beta_{k-1}^{(k)} p_{k-1}) = 0,$$

$$(r_k, p_k) = \left(r_k, r_k - \sum_{j=0}^{k-1} \beta_j^{(k-1)} p_j \right) = \|r_k\|^2, \quad \left(\|x\| \equiv \sqrt{(x, x)} \right)$$

$$\therefore \beta_k^{(k)} = -\frac{\|r_{k+1}\|^2}{\|r_k\|^2}, \quad p_{k+1} = r_{k+1} - \beta_k^{(k)} p_k.$$

算法 ~ 共役勾配法 ~

$$x_0 \text{ を適当にとり } r_0 = p_0 = b - A x_0;$$

$$k = 0, 1, 2, \dots \text{ に対して}$$

$$\left[\begin{array}{l} \alpha_k = \frac{\|r_k\|^2}{(p_k, A p_k)}; \\ x_{k+1} = x_k + \alpha_k p_k; \\ r_{k+1} = r_k - \alpha_k A p_k; \\ \|r_{k+1}\| \leq \varepsilon \|b\| \text{ ならば 終了}; \quad (\varepsilon \text{ を小さい正数}) \\ \beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}; \\ p_{k+1} = r_{k+1} + \beta_k p_k; \end{array} \right.$$

< CG法の改良 ~ 前処理 ~ >

CG法 --- M.R. Hestenes & E. Stiefel:
 Methods of conjugate gradients for solving linear systems,
 Journal of Research of the National Bureau of Standards,
 Vol. 49 (1952), 409-436.

しかし、CG法の収束速度は問題に依存する、長期間
 収束しないことがある。

↓
 1980年: 前処理 を併用するとCG法の“使用”が速くなる。
 → CG法の復活

非対称行列への拡張、安定化に関する応用 etc.

前処理 (preconditioning)

CG法の、単位行列に近い係数行列に対する収束速度。

* 定理:
$$\frac{f(x_k) - f(x^*)}{f(x_0) - f(x^*)} \leq \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k}$$

ここで、 $\kappa = \frac{A \text{ の最大固有値}}{A \text{ の最小固有値}}$

(A が正定値対称行列ならば、 A の固有値 > 0 に注意)

(参考) 杉原正昭・室田一雄「線形計算の数理」,
 岩波出版「マセコ」シリーズ, 2016年。

さらに、

何らかの法で近似的に $A \approx C C^T$ と分解する (C : 正則行列)。

↓

$$A x = b \Leftrightarrow \underbrace{(C^{-1} A C^{-T})}_{\text{単位行列 } I \text{ に近い}} C^T x = C^{-1} b \quad (C^{-T} = (C^T)^{-1} = (C^{-1})^T)$$

単位行列 I に近い。

$Ax = b$ の代わりに $\tilde{A}\tilde{x} = \tilde{b}$ ($\tilde{A} = C^{-1}AC^{-T}$, $\tilde{b} = C^{-1}b$, $\tilde{x} = C^T x$) を解くのは, CG法の収束が速くなることを期待できる.
 (実際, 速くなる) ... 適切な前処理.

算法 ~ 前処理付き共役勾配法 (PCG法) ~

x_0 を適当に選ぶ;

$r_0 = b - Ax_0$;

$p_0 = (CC^T)^{-1} r_0$;

$k = 0, 1, 2, \dots$ まで

$$\left[\begin{array}{l} \alpha_k = \frac{(r_k, (CC^T)^{-1} r_k)}{(p_k, A p_k)} ; \\ x_{k+1} = x_k + \alpha_k p_k ; \\ r_{k+1} = r_k - \alpha_k A p_k ; \\ \|r_{k+1}\| \leq \epsilon \|b\| \text{ まで, } r_k \text{ が小さくなった;} \\ \beta_k = \frac{(r_{k+1}, (CC^T)^{-1} r_{k+1})}{(r_k, (CC^T)^{-1} r_k)} ; \\ p_{k+1} = (CC^T)^{-1} r_{k+1} + \beta_k p_k ; \end{array} \right.$$

(注意) $y = (CC^T)^{-1} x$ の計算は連立一次方程式

$$\begin{cases} Cy = x \\ C^T y = z \end{cases}$$

を解く必要がある (逆行列の計算を避ける) .

PCG法の算法の導出: $\tilde{A}\tilde{x} = \tilde{b}$ はCG法を適用する次のようにする:

\tilde{x}_0 を適当に選ぶ;

$\tilde{r}_0 = \tilde{b} - \tilde{A}\tilde{x}_0$;

$k = 0, 1, 2, \dots$ まで

$$\left[\begin{array}{l} \tilde{\alpha}_k = \|\tilde{r}_k\|^2 / (\tilde{p}_k, \tilde{A}\tilde{p}_k) ; \\ \tilde{x}_{k+1} = \tilde{x}_k + \tilde{\alpha}_k \tilde{p}_k ; \\ \tilde{r}_{k+1} = \tilde{r}_k - \tilde{\alpha}_k \tilde{A}\tilde{p}_k ; \\ \tilde{\beta}_k = \|\tilde{r}_{k+1}\|^2 / \|\tilde{r}_k\|^2 ; \\ \tilde{p}_{k+1} = \tilde{r}_{k+1} + \tilde{\beta}_k \tilde{p}_k ; \end{array} \right.$$

$$\tilde{x} = C^T \alpha, \quad \tilde{r}_k = C^T r_k, \quad \tilde{p}_k = C^T p_k \quad \text{と } \tilde{A} = C^T A C^{-1}$$

$$\tilde{r}_k = \tilde{b} - \tilde{A} \tilde{x}_k = C^{-1} b - (C^{-1} A C^{-T}) (C^T \alpha_k)$$

$$= C^{-1} (b - A \alpha_k) = C^{-1} r_k$$

よって $\tilde{r}_k = C^{-1} r_k$ とおくと $\tilde{A} \tilde{p}_k = \tilde{r}_k$ とおくと

$$(\tilde{p}_k, \tilde{A} \tilde{p}_k) = (C^T p_k, (C^{-1} A C^{-T}) (C^T p_k)) = (C^T p_k, C^{-1} A p_k)$$

$$= (p_k, C C^{-1} A p_k) = (p_k, A p_k),$$

$$\|\tilde{r}_k\|^2 = (\tilde{r}_k, \tilde{r}_k) = (C^{-1} r_k, C^{-1} r_k) = (r_k, C^{-T} C^{-1} r_k)$$

$$= (r_k, (C C^T)^{-1} r_k)$$

よって, $\tilde{\alpha}_k = \frac{(r_k, (C C^T)^{-1} r_k)}{(p_k, A p_k)}, \quad \tilde{\beta}_k = \frac{(r_{k+1}, (C C^T)^{-1} r_{k+1})}{(r_k, (C C^T)^{-1} r_k)}$

α_k, r_k, p_k の更新式は

$$\tilde{\alpha}_{k+1} = C^T \alpha_{k+1} = \tilde{\alpha}_k + \tilde{\alpha}_k \tilde{p}_k = C^T (\alpha_k + \tilde{\alpha}_k p_k),$$

$$\rightarrow \alpha_{k+1} = \alpha_k + \tilde{\alpha}_k p_k,$$

$$\tilde{r}_{k+1} = C^{-1} r_{k+1} = \tilde{r}_k - \tilde{\alpha}_k \tilde{A} \tilde{p}_k = C^{-1} r_k - \tilde{\alpha}_k (C^{-1} A C^{-T}) C^T p_k$$

$$= C^{-1} (r_k - \tilde{\alpha}_k A p_k)$$

$$\rightarrow r_{k+1} = r_k - \tilde{\alpha}_k A p_k,$$

$$\tilde{p}_{k+1} = C^T p_{k+1} = \tilde{r}_{k+1} + \tilde{\beta}_k \tilde{p}_k = C^{-1} r_{k+1} + \tilde{\beta}_k C^T p_k$$

$$= C^{-1} (r_{k+1} + \tilde{\beta}_k C C^T p_k)$$

$$\rightarrow p_{k+1} = (C C^T)^{-1} r_{k+1} + \tilde{\beta}_k p_k$$

と更新される。



不完全 Cholesky 分解

前処理の例.

対称行列 A の LU 分解を施す
 \rightarrow Cholesky 分解

$$A = LDL^T, \quad L = \begin{bmatrix} l_{11} & & 0 \\ & \ddots & \\ l_{m1} & \dots & l_{mn} \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{bmatrix}$$

不完全 Cholesky 分解 ... Cholesky 分解を近似的に行う

$$A \approx LDL^T$$

具体的な u, v

- $l_{ii} d_i = 1 \quad (i = 1, \dots, n)$ とする.
- $a_{ij} = 0$ ならば (i, j) かつ u, v $l_{ij} = 0$ とする

↓

$$\boxed{\begin{array}{l} k = 1, 2, \dots, n \text{ に対して} \\ \left[\begin{array}{l} i = 1, 2, \dots, k-1 \text{ に対して} \\ \left[\begin{array}{l} l_{ki} = a_{ki} - \sum_{j=1}^{i-1} l_{kj} l_{ij} d_j; \\ d_k = \left(a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2 d_j \right)^{-1} (= l_{kk}^{-1}); \end{array} \right. \end{array} \right. \end{array}$$

$y_1 = (LDL^T)^{-1} x$ の計算 ... $LDL^T y_1 = x$ を前進/後退代入で解く.

ICCG法 ... 前処理として不完全 Cholesky 分解を用いた CG 法.

(I ... "incomplete" の頭文字)

(参考) 森正武「Fontnam 77 対称計算機の「ラミエ」増補版」
 岩波書店, 1987年.