共役勾配法

辻 裕太

2017/5/4

1 CG法

 $A \in M_n(\mathbb{R}), \quad x, f \in \mathbb{R}^n$ として連立一次方程式 Ax = f を解く。前回は線形反復法と呼ばれる手法について解説した。今回は CG 法と呼ばれる手法について解説する。ここでは CG 法は反復法の一種として扱っているが、実は有限回で厳密解が得られる (!) という性質から直接法として扱われることもある。

1.1 変分原理

$$A \in M_n(\mathbb{R}), \quad x, f \in \mathbb{R}^n$$

とする。A, f は既知、x は未知として連立一次方程式

$$Ax = f$$

を考える。

 $F:\mathbb{R}^n\to\mathbb{R}$

を

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

として定める。 A が対称正定値行列ならば、次が成立する。

Fact 1.1. F(x) が最小値を取る $x \in \mathbb{R}^n$ が一意に存在して、Ax = f を満たす。

つまりは元の方程式をある関数の最小化問題にすり替えるわけである。CG 法ではアイデアとしてこれを使うが、収束の議論の際には使わないので証明は省略する。ただしこのような定式化 (変分問題という) は、有限要素法の根幹をなす部分なので、いずれするであろう有限要素法の話の時に、一般化された形で証明する。

1.2 CG 法の原理

この節では、A は対称正定値であるとする。前節の事実から、F(x) を最小化する x を求めればよい。そこで適当な x から初めて新しい x を次々に定めて、F(x) を可能な限り小さくしていくことを考える。まずは次の定理を示そう。

Theorem 1.2.

$$x, u \in \mathbb{R}^n, \quad u \neq 0$$

とする。 $f: \mathbb{R} \to \mathbb{R}$ を

$$g(\alpha) := F(x + \alpha u)$$

で定めると、g が最小値を取る α がただ一つ存在し、

$$\alpha = \frac{(f - Ax, u)}{(Au, u)}$$

となる。

Proof.

$$\begin{split} g(\alpha) &= F(x + \alpha u) \\ &= \frac{1}{2}(A(x + \alpha u), x + \alpha u) - (f, x + \alpha u) \\ &= \frac{1}{2}(Ax, x) + \frac{1}{2}(Ax, \alpha u) + \frac{1}{2}(A\alpha u, x) \\ &+ \frac{1}{2}(A\alpha u, x) + \frac{1}{2}(A\alpha u, \alpha u) - (f, x) - (f, \alpha u) \\ &= \frac{\alpha^2}{2}(Au, u) + \alpha(Ax, u) - \alpha(f, u) + \frac{1}{2}(Ax, x) - (f, x) \end{split}$$

ただし途中で A が対称であることを用いている。この式は α の二次式であり、A は正定値だから二次の係数 $\frac{1}{2}(Au,u)$ は正である。したがって最小値は存在し、平方完成することで

$$\alpha = \frac{(f - Ax, u)}{(Au, u)}$$

となる。

この定理は x,u を与えたときに、より F を小さくするための $x+\alpha u$ について、 α として適当なものを具体的に書いている。そこでどのように u を取るのかが問題になる。F が微分可能ならば $F(x+y)-F(x)=(\nabla F(x),y)+o(|y|)$ と書けるので、y を $\nabla F(x)$ と平行になるように取れば変化量を大きく出来そうである。実際 $y=-\nabla F(x)$ とする手法も存在するが、実はより効率の良い方法が存在する。その準備として Krylov 部分空間を定義する。最初はなぜこのようなものを考えるかわかり辛い (と思った) が、これこそが CG 法の原理でもっとも重要なのである。

1.3 Krylov 部分空間

この節では、

$$A \in GL_n(\mathbb{R}), \quad f \in \mathbb{R}^n, \quad f \neq 0$$

とする。

Definition 1.3. $k \in \mathbb{N}$ に対して,Krylov 部分空間 $S_A^k(f)$ を

$$S_A^k(f) := Span\{f, Af, \dots, A^{k-1}f\}$$

として定める。ただし、 $Span\{f,Af,\dots,A^{k-1}f\}$ は $\{f,Af,\dots,A^{k-1}f\}$ の張る部分空間を表す。 定義から次が従うことは明らかであろう。

$$S_A^k(f) \subset S_A^{k+1}(f)$$

$$1 \le \dim S_A^k(f) \le n$$

これから、 $S_A^k(f)$ の次元を d_k とすると、 d_k は単調増加で上に有界だから

$$\exists m \in \mathbb{N}, \quad d_m = d_{m+1}$$

となる。さらにこの m について次の定理が成立する。

Theorem 1.4.

$$k \ge m \Rightarrow d_k = d_m$$

Proof. $k \geq m$ の時、 $a_i \in \mathbb{C}(i=0,\ldots,m-1)$ があって

$$A^k f = \sum_{i=0}^{m-1} a_i A^i f$$

と書けることを数学的帰納法で証明する。定理はこれから直ちに従う。まず、

$$d_m = d_{m+1}$$

であるから、 $a_i \in \mathbb{C}(i=0,\ldots,m-1)$ があって

$$A^{m}f = \sum_{i=0}^{m-1} a_{i}A^{i}f$$

と書ける。よって k=m では成立する。 $m \le k \le n$ となる全ての k で成立すると仮定すると、仮定から

$$A^{n+1}(f) = A(A^n f)$$

$$= A(\sum_{i=0}^{n-1} a_i A^i f)$$

$$= \sum_{i=1}^n a_{i-1} A^i f$$

となるが、再び仮定から $A^if(i=1,\dots,n)$ は $f,Af,\dots,A^{m-1}f$ の一次結合で書けるので k=n+1 でも成立する。よって示された。

さらに次が成立する。

Theorem 1.5. ある $m \in \mathbb{N}$ が存在して、次の条件を満たす。

$$1 \le k \le m \Rightarrow d_k = k$$

$$k \ge m \Rightarrow d_k = m$$

Proof. Theorem 1.4 の m のうち、最小となるものをとればよい。

実は解は Krylov 部分空間の元になる。それを示そう。

Theorem 1.6. Theorem 1.5 の m と、 $x \in \mathbb{R}^n$ を未知とする方程式 Ax = f の解 $x = A^{-1}f$ について、次が成立する。

$$x \in S_A^m(f)$$

$$k < m \Rightarrow x \notin S_A^k(f)$$

Proof. $d_m = d_{m+1}$ より $A^m f \in S^m_A(f)$ だから、 $\{a_i\}_{i=1}^{m-1} \subset \mathbb{C}$ があって

$$A^m f = \sum_{i=0}^{m-1} a_i A^i f$$

と書ける。両辺に A^{-1} を掛けて

$$a_0 A^{-1} f = -A^{m-1} f + \sum_{i=1}^{m-1} a_i A^{i-1} f$$

を得る。 $a_0=0$ とすると $A^{m-1}f\in S_A^{m-1}(f)$ となって $d_m=d_{m-1}$ となるが、これは m の定義に反する。 したがって $a_0\neq 0$ であり、

$$A^{-1}f = -\frac{1}{a_0}A^{m-1}f + \frac{1}{a_0}\sum_{i=1}^{m-1}a_iA^{i-1}f$$

となる。したがって

$$x \in S_A^m(f)$$

であり、特に $A^{m-1}f$ の係数はゼロでないから、

$$k < m \Rightarrow x \notin S_A^k(f)$$

が成立する。

この定理から、真の解が存在する空間に有限回基底を足して部分空間を拡張することでたどり着けることが 分かる。

1.4 計算原理

前述の Krylov 部分空間の基底として、具体的にどのようなものを取るのが良いかを考える。

Definition 1.7.

$$x, y \in \mathbb{R}^n$$

に対し、 $(\cdot,\cdot)_A:\mathbb{R}^n imes\mathbb{R}^n o\mathbb{R}$ を

$$(x,y)_A := (Ax,y)$$

として定義する。ただし上式右辺は \mathbb{R}^n の通常の内積である。

これが再び \mathbb{R}^n の内積になることは明らかであろう。

Definition 1.8.

$$x, y \in \mathbb{R}^n$$

に対し、x,y が共役であるとは、

$$(x,y)_A = (Ax,y) = 0$$

となることを言う。

Krylov 部分空間 $S_A^m(f)$ の基底として (ϕ_1, \ldots, ϕ_m) で

$$(\phi_i, \phi_j)_A = 0 \quad (i \neq j, \quad 1 \leq i, j \leq n)$$

となるもの (共役な基底) を取る。すると $x=A^{-1}f$ は、直交基底の性質から

$$x = \sum_{i=1}^{m} (x, \phi_i)_A \frac{\phi_i}{\|\phi_i\|_A^2}$$

と書けるが、 $(x,\phi_i)_A = (f,\phi_i)$ だから

$$x = \sum_{i=1}^{m} (f, \phi_i) \frac{\phi_i}{\|\phi_i\|_A^2}$$

となる。なんと右辺からxが消えてしまう(!)そこで

$$x^{(0)} = 0$$

$$x^{(k)} = \sum_{i=1}^{k} \frac{(f, \phi_i)}{\|\phi_i\|_A^2} \phi_i \quad (1 \le k \le m)$$

として $x^{(k)}$ を定めると、

$$x^{(m)} = A^{-1}f$$

であり、 $1 \le k \le m$ に対して

$$x^{(k)} = x^{(k-1)} + \frac{(f, \phi_k)}{\|\phi_k\|_A^2} \phi_k$$

$$(x^{(k-1)}, \phi_k)_A = 0$$

が成立する。この関係式が成立するのが大切である。元々は、

$$\alpha = \frac{(f - Ax, u)}{\|u\|_A^2}$$

として新しい x を

$$x + \alpha u$$

で定めるのがアイデアだったから、 $u = \phi_k$ (共役な基底)、x を x_k として、

$$x_0 = 0, \quad x_k = x_{k-1} + \alpha_{k-1}\phi_k$$

$$\alpha_{k-1} = \frac{(f - Ax_{k-1}, \phi_k)}{\|\phi_k\|_A^2}$$

として反復する。 $(x_{k-1}, \phi_k)_A = 0$ が成立するから、

$$\alpha_{k-1} = \frac{(f - Ax^{(k-1)}, \phi_k)}{\|\phi_k\|_A^2}$$
$$= \frac{(f, \phi_k)}{\|\phi_k\|_A^2}$$

となり、なんと x_k は先に定めた $x^{(k)}$ と一致する。 $x^{(m)}$ は、厳密解そのものであった。従って、これは m 回の反復で厳密解に到達するのである (!) これは共役な基底を取れれば成立する議論である。もちろん今考えている空間は有限次元であり、 $(,)_A$ が内積であることから、適当なベクトルから出発してシュミットの直交化法を用いれば共役な基底が取れることは明らかであるが、数値計算をする上では具体的にどう定めるかが重要であるから、次節でそれを行おう。

1.5 計算手順

$$x^{(0)} \in \mathbb{R}^n$$

を出発ベクトルとする。これは一般には0ではないから、上の議論はそのままでは適用できない。しかし、

$$f^* = f - Ax^{(0)}$$

とすれば、

$$Ax = f^*$$

の初期ベクトルを 0 に取ったのと同じである。そこで以降はこの方程式について議論する。この方程式に Theorem 1.5 を適用したものを m とする。 $\phi^{(1)}=f^*$ とすると、明らかに $\phi^{(1)}\in S^1_Af^*$ である。

$$r^{(0)} = f^*$$

として残差ベクトルを定めよう。これが 0 に等しいならば $x^{(0)}$ は解であるから、ここで反復を打ち切ればよい。0 でないならば、 $r^{(0)} \in S^1_A f^*$ である。

$$\alpha^{(1)} = \frac{(r^{(0)}, \phi^{(1)})}{\|\phi^{(1)}\|_A^2}$$

として

$$x^{(1)} = x^{(0)} + \alpha^{(1)}\phi^{(1)}$$

と定める。 $r^{(0)} = \phi^{(1)}$ より、 $\alpha^{(1)} \neq 0$ である。残差を

$$r^{(1)} = f - Ax^{(1)}$$

とする。これが0なら $x^{(1)}$ を解として反復を終了する。そうでないとき、 $x^{(1)}$ を代入して、

$$r^{(1)} = f^* - \alpha^{(1)} A \phi^{(1)}$$
$$= r^{(0)} - \alpha^{(1)} A \phi^{(1)}$$

となる。 $\alpha^{(1)} \neq 0$ 、 $\phi^{(1)} = f^*$ から

$$r^{(1)} \in S_A^2(f^*) - S_A^1(f^*)$$

となる。重要なのは、共役な基底として取りたい $\phi^{(k)}$ の決定である。 $\phi^{(2)}$ は、この時点ではシュミットと同様にして

$$\phi^{(2)} = r^{(1)} - \frac{(\phi^{(1)}, r^{(1)})_A}{\|\phi^{(1)}\|_A^2} \phi^{(1)}$$

として定めよう。以下同様に

$$\alpha^{(k)} = \frac{(r^{(k-1)}, \phi^{(k)})}{\|\phi^{(k)}\|_A^2}$$

$$x^{(k)} = x^{(k-1)} + \alpha^{(k)}\phi^{(k)}$$

$$r^{(k)} = r^{(k-1)} - \alpha^{(k)} A \phi^{(k)}$$

とする。しかし $\phi^{(k)}$ だけは

$$\phi^{(k+1)} = r^{(k)} - \frac{(\phi^{(k)}, r^{(k)})_A}{\|\phi^{(k)}\|_A^2} \phi^{(k)}$$

として定める。これはシュミットの方法とは一致しない。右辺第二項が和になっていない分、こちらの方が計算回数が少なくて済む。ただし問題なのは、こうして定めることでも共役な基底になることが明らかではないことである。次節では実際に共役な基底であることを示す。具体的な計算手順としてまとめておこう。

$$\phi^{(1)} = f^* = f - Ax^{(0)} = r^{(0)}$$

として、 $k \ge 1$ の時は

$$\alpha^{(k)} = \frac{(r^{(k-1)}, \phi^{(k)})}{\|\phi^{(k)}\|_A^2}$$

$$x^{(k)} = x^{(k-1)} + \alpha^{(k)} \phi^{(k)}$$

$$\beta^{(k)} = \frac{(\phi^{(k)}, r^{(k)})_A}{\|\phi^{(k)}\|_A^2}$$

$$\phi^{(k+1)} = r^{(k)} - \beta^{(k)}\phi^{(k)}$$

とする。

1.6 収束

前節で CG 法の具体的な手順を書いた。この節では確かにこの手順で収束する (有限回で厳密解にたどり着く) ことを示そう。とは言っても主要な議論はこれまでの段階で終わっていて、共役な基底が取れてさえいれば収束することはもう述べてある。残るのは CG 法が上手く定義できていることと、確かに $\phi^{(k)}$ が共役な基底を定めているということの証明である。途中で $r^{(k)}$ の直交性なども経由するので、少し長くなる。 CG 法の手順において、 $m\leq 2$ の時は $\phi^{(k)}$ の定義の仕方はシュミットの直交化法と同じだったので、解にたどり着くことは明らか。従って、以下では m>2 とする。

Theorem 1.9. $x, f \in \mathbb{R}^n$ 、 $A \in M_n(\mathbb{R})$ は対称正定値行列とする。

$$Ax = f$$

について、前節の手順で $\alpha^{(k)}, r^{(k)}, \phi^{(k)}$ を定める。1 < k < mを満たす任意のkに対して、次が成立する。

$$S_A^k(f^*) = \operatorname{Span}\{r^{(i-1)}\}_{i=1}^k = \operatorname{Span}\{\phi^{(i)}\}_{i=1}^k$$

$$\alpha^{(k)} \neq 0$$

$$(r^{(i-1)}, r^{(j-1)}) = 0, \quad (\phi^{(i)}, \phi^{(j)})_A = 0 \quad (i \neq j, \quad 1 \leq i, j \leq k)$$

ただし、最後は $k \ge 2$ の時のみ考える。

Proof. 帰納法で示す。k=1,2 の時に成立することは定め方から分かる $(r^{(k)}$ の直交性は $\alpha^{(k)}$ の定義から従う)。 $1 \leq k \leq n-1$ の時に成立すると仮定して、k=n の時も成立することを示す。 $S_A^k(f^*)$ の次元は k < m の時は単調増加であるから、この仮定の下で

$$r^{(i-1)} \neq 0, \quad \phi^{(i)} \neq 0 \quad (i = 1, \dots, n-1)$$

となる事に注意する。まず $r^{(k)}$ の直交性について、仮定から $(r^{(i-1)}, r^{(j-1)}) = 0$ $(i \neq j, 1 \leq i, j \leq n-1)$ だから、

$$(r^{(n-1)}, r^{(j)}) = 0 \quad (j = 1, \dots, n-2)$$

を示せば、k=n についても成立する。これを示す。仮定から

$$\mathrm{Span}\{r^{(i-1)}\}_{i=1}^{n-1}=\mathrm{Span}\{\phi^{(i)}\}_{i=1}^{n-1}$$

だから、

$$(r^{(n-2)}, \phi^{(j)}) \quad (1 \le j \le n-2)$$

について、 $\phi^{(j)}$ は $r^{(i-1)}(i=1,\dots,n-1)$ の線形結合で書ける。これと仮定から

$$(r^{(i-1)}, r^{(j-1)}) = 0 \quad (i \neq j, 1 \leq i, j \leq n-1)$$

なので、

$$(r^{(n-2)}, \phi^{(j)}) = 0 \quad (1 \le j \le n-2)$$

となる。また、 $r^{(n-1)} = r^{(n-2)} - \alpha^{(n-1)} A \phi^{(n-1)}$ より

$$(r^{(n-1)}, \phi^{(n-1)}) = (r^{(n-2)}, \phi^{(n-1)}) - \alpha^{(n-1)}(\phi^{(n-1)}, \phi^{(n-1)})_A$$

となるが、

$$\alpha^{(n-1)} = \frac{(r^{(n-2)}, \phi^{(n-1)})}{\|\phi^{(n-1)}\|_A^2}$$

だから、

$$(r^{(n-1)}, \phi^{(n-1)}) = 0$$

となる。したがって、

$$(r^{(n-1)}, \phi^{(j)}) = 0 \quad (1 \le j \le n-1)$$

となる。再び定理の仮定から

$$\mathrm{Span}\{r^{(i-1)}\}_{i=1}^k = \mathrm{Span}\{\phi^{(i)}\}_{i=1}^k \quad (1 \le k \le n-1)$$

となるので、 $\phi^{(j)}$ は $r^{(i-1)}(j=1,\ldots,j)$ の線形結合で書け、特に $r^{(j-1)}$ の係数は 0 でない。そこで j=1 から順に考えることで、

$$(r^{(n-1)}, r^{(j)})$$
 $(j = 1, \dots, n-2)$

となる。よって $r^{(k)}$ の直交性は k=n でも成立する。次に仮定の下で

$$S_A^n(f^*) = \text{Span}\{r^{(i-1)}\}_{i=1}^n = \text{Span}\{\phi^{(i)}\}_{i=1}^n$$

となる事を示す。まず

$$\mathrm{Span}\{r^{(i-1)}\}_{i=1}^n = \mathrm{Span}\{\phi^{(i)}\}_{i=1}^n$$

を示す。

$$\phi^{(n)} = r^{(n-1)} + \beta^{(k)}\phi^{(n-1)}$$

だから、仮定より $\phi^{(n-1)} \in \operatorname{Span}\{r^{(i-1)}\}_{i=1}^{n-1}$ なので、

$$\phi^{(n)} \in \operatorname{Span}\{r^{(i-1)}\}_{i=1}^n$$

となり、

$$\operatorname{Span}\{r^{(i-1)}\}_{i=1}^n \supset \operatorname{Span}\{\phi^{(i)}\}_{i=1}^n$$

が成り立つ。反対側は

$$r^{(n-1)} = \beta^{(k)}\phi^{(n-1)} - \phi^{(n)}$$

から成立する。よって示された。次に

$$S_A^n(f^*) = \operatorname{Span}\{r^{(i-1)}\}_{i=1}^n$$

を示す。仮定から $\phi^{(n-1)} \in S^{n-1}_A(f^*)$ なので、

$$A\phi^{(n-1)} \in S_A^n(f^*)$$

となる。

$$r^{(n-1)} = r^{(n-2)} - \alpha^{(n-1)} A \phi^{(n-1)}$$

だから、仮定より $r^{(n-2)} \in S^{n-1}_A(f^*)$ なので

$$r^{(n-1)} \in S_A^n(f^*)$$

を得る。したがって

$$S_A^n(f^*) \supset \operatorname{Span}\{r^{(i-1)}\}_{i=1}^n$$

となる。仮定より

$$A^{n-2}f^* \in \text{Span}\{\phi^{(i)}\}_{i=1}^{n-1}$$

だから、

$$A^{n-1}f^* \in \text{Span}\{A\phi^{(i)}\}_{i=1}^{n-1}$$

となる。

$$r^{(i)} = r^{(i-1)} - \alpha^{(i)} A \phi^{(i)} \quad (i = 1, \dots, n-1)$$

より、仮定から $\alpha^{(i)} \neq 0$ だから、

$$A\phi^{(i)} \in \text{Span}\{r^{(j-1)}\}_{j=1}^{i+1} \quad (i=1,\dots,n-1)$$

である。したがって

$$A^{n-1}f^* \in \text{Span}\{r^{(i-1)}\}_{i=1}^n$$

であり、

$$S_A^n(f^*) \subset \text{Span}\{r^{(i-1)}\}_{i=1}^n$$

となる。よって

$$S_A^n(f^*) = \text{Span}\{r^{(i-1)}\}_{i=1}^n$$

となり、示された。次に、仮定の下で

$$\alpha^{(n)} \neq 0$$

となる事を示す。 $\phi^{(n)}$ の漸化式から、

ははいり、
$$\alpha^{(n)} = \frac{(r^{(n-1)}, \phi^{(n)})}{\|\phi^{(n)}\|_A^2}$$
$$= \frac{(r^{(n-1)}, r^{(n-1)}) - \beta^{(n-1)}(r^{(n-1)}, \phi^{(n-1)})}{\|\phi^{(n)}\|_A^2}$$

となるが、残差についての議論の途中で $(r^{(n-1)},\phi^{(n-1)})=0$ であることが分かっており、また先に示した

$$S_A^n(f^*) = \operatorname{Span}\{r^{(i-1)}\}_{i=1}^n = \operatorname{Span}\{\phi^{(i)}\}_{i=1}^n$$

から、 $r^{(n-1)} \neq 0$, $\phi^{(n)} \neq 0$ である。したがって

$$\alpha^{(n)} \neq 0$$

であり、示された。最後に、 $\phi^{(k)}$ の共役性が k=n でも成立することを示す。 $r^{(k)}$ の時と同様に、

$$(\phi^{(n)}, \phi^{(j)})_A = 0 \quad (j = 1, \dots, n-1)$$

を示せばよい。残差の漸化式

$$r^{(j)} = r^{(j-1)} - \alpha^{(j)} A \phi^{(j)} \quad (j = 1, \dots, m)$$

について、 $1 \le j \le n-1$ の時は仮定から $\alpha^{(j)} \ne 0$ であり、割ることで

$$A\phi^{(j)} \in \text{Span}\{r^{(i-1)}\}_{i=1}^{j+1} \quad (1 \le j \le n-1)$$

となる。このことと $r^{(k)}$ の直交性から、

$$(r^{(n-1)}, \phi^j)_A = (r^{(n-1)}, A\phi^j) = 0 \quad (j = 1, \dots, n-2)$$

となる。したがって、 $j=1,\ldots,n-2$ で

$$(\phi^{(n)}, \phi^{(j)})_A = (r^{(n-1)} + \beta^{(n-1)}\phi^{(n-1)}, \phi^{(j)})$$
$$= (r^{(n-1)}, \phi^{(j)}) + \beta^{(n-1)}(\phi^{(n-1), \phi^{(j)}})$$

となるが、仮定より $(\phi^{(n-1),\phi^{(j)}})=0$ であることと、先の議論から

$$(\phi^{(n)}, \phi^{(j)})_A = 0 \quad (j = 1, \dots, n-2)$$

となる。 $(\phi^{(n)},\phi^{(n-1)})_A=0$ となることは $\phi^{(n)}$ をこうなるよう定めたから成立する。以上により、

$$(\phi^{(n)}, \phi^{(j)})_A = 0 \quad (j = 1, \dots, n-1)$$

となる。よって示された。これでk=nでも成立し、証明が終わった。