

# 大規模線形方程式を解くためのクリロフ 部分空間法の前処理

東京工業大学 社会理工学研究科 中田 和秀

## 1 はじめに

$A \in \mathbb{C}^{m \times m}$ ,  $b \in \mathbb{C}^m$  が与えられているとき、線形方程式系

$$Ax = b \quad (1)$$

の解  $x \in \mathbb{C}^m$  を求めたい。線形方程式系 (1) の解法として、LU 分解などの直接法が良く知られている。しかし直接法では、この線形方程式系 (1) のサイズ  $m$  に対し、一般に  $O(m^3)$  flops の計算量と係数行列  $A$  の記憶が必要となる。このため、大規模な線形方程式系を解くことは困難である。たとえ係数行列  $A$  が疎行列であっても、LU 分解などの直接法では、計算途中で fill-in が起こりその疎性が崩れる事が多い。一方、クリロフ部分空間法では係数行列を直接に操作することは無く、ベクトルに対する線形変換にのみ利用するため、係数行列の疎性の維持が容易である。また、クリロフ部分空間法の収束が速ければ高速に近似解を得る可能性がある。このため、本論文ではクリロフ部分空間法により線形方程式系 (1) を解くことを試みる。2 章では、4 つのクリロフ部分空間法と前処理の一般的な枠組みについて述べる。3 章では、幾つかの前処理行列について説明を行う。4 章では半正定値計画問題を例に取り、係数行列が密である特殊な線形方程式系に対して、クリロフ部分空間法が有効であることを示す。

## 2 クリロフ部分空間法とその前処理

クリロフ部分空間  $\mathcal{K}_n(A; b)$  を、

$$\mathcal{K}_n(A; b) = \text{span}\{b, Ab, A^2b, \dots, A^{n-1}b\}$$

で定義する。クリロフ部分空間法は非定常な反復解法であり、 $n$  反復目の反復点  $x_n$  がクリロフ部分空間  $\mathcal{K}_n(A; b)$  の中をなんらかの意味で探索する方法である。(本論文では、記述を簡略化するため、クリロフ部分空間法の初期点をととして  $x_0 = 0$  を仮定する。) これまで、多くのクリロフ部分空間法が提案されてきたが、歴史的には係数行列が正定値行列である線形方程式系に対する解法として、Hestenes と Stiefel が共役勾配法 (Conjugate Gradient method: CG) を発表したことに始まる [6]。その後、同じく係数行列が正定値な線形方程式系に対する解法である共役残差法 (Conjugate Residual method: CR) や、一般の線形方程式系に対する解法である双共役勾配法 (Biconjugate Gradient

method: Bi-CG)、2乗共役勾配法 (Conjugate Gradient Squared method: CGS)、安定化双共役勾配法 (Biconjugate Gradient Stabilized method: Bi-CGSTAB)、疑似最小残差法 (Quasi-Minimal Residual method: QMR)、一般化共役残差法 (Generalized Conjugate Residual method: GCR)、リスタート版一般化共役残差法 (Restarted Generalized Conjugate Residual method: GCR(m))、一般化最小残差法 (Generalized Minimal Residual method: GMRES)、リスタート版一般化最小残差法 (Restarted Generalized Minimal Residual method: GMRES(m)) などのクリロフ部分空間法が提案されている [2, 3, 5]。

クリロフ部分空間法の収束性は係数行列  $A$  の固有値の分布に依存し、特に係数行列  $A$  の固有値が密集している場合には高速に収束することが知られている。このため、線形方程式系 (1) を解く代わりに、正則行列  $M_1, M_2 \in C^{m \times m}$  に対し、同値な線形方程式系

$$(M_1 A M_2)(M_2^{-1} x) = (M_1 b) \quad (2)$$

を考え、この線形方程式系をクリロフ部分空間法で解くことが有効である。このとき、 $M_1 A M_2$  の固有値が密集していれば、新しいクリロフ部分空間法は高速に収束することが期待できる。この技術を前処理と呼び、線形方程式系 (2) をクリロフ部分空間法で解くことを前処理付きクリロフ部分空間法という。元の線形方程式系 (1) の係数行列  $A$  が正定値である場合、その正定値性を壊さない前処理が求められる。そのような場合、線形方程式系 (2) の代わりに、正則行列  $M \in C^{m \times m}$  に対し、

$$(M A M^*)(M^{-*} x) = (M b) \quad (3)$$

という線形方程式系を考え、これに共役勾配法や共役残差法を適用することが大変有効である。

本章では以下、典型的な4つのクリロフ部分空間法を例に取り、そららに対し前処理を適用することにより、前処理付きのクリロフ部分空間法の概要を説明する。

## 2.1 共役勾配法

本節では、クリロフ部分空間法の中で最も基本的かつ有名な方法である共役勾配法について述べる。共役勾配法は、係数行列が正定値行列である線形方程式系を解くクリロフ部分空間法である。前節で述べたように、クリロフ部分空間法は  $n$  反復目の反復点  $x_n$  がクリロフ部分空間  $\mathcal{K}_n(A; b)$  を探索する反復解法であるが、特に共役勾配法では反復点  $x_n$  が

$$\min_{x \in \mathcal{K}_n(A; b)} \|b - Ax\|_{A^{-1}} \quad (4)$$

の解となっていることが知られている。この性質は、理論的には共役勾配法が高々  $m$  回の反復で線形方程式系の厳密解を得ることを保証する。しかし、一般にクリロフ部分空間法の計算は誤差が生じやすいことが知られており、共役勾配法において  $m$  回の反復での終了は期待できない場合も多い。アルゴリズム1で線形方程式系 (1) を解く共役勾配法を示している。アルゴリズム中の  $\alpha, \beta$  には、理論的には同じ値を導く幾つかの計算方法が存在することを付記しておく。係数行列  $A$  はベクトルの線形変換  $A p_n$  にのみ使われている。ここで  $\#(A)$  を行列  $A$  の非ゼロ要素の数とすると、共役勾配法の1反復当りの計算量は、 $2 \#(A) + \mathcal{O}(m)$  flops となる。

アルゴリズム 1:  
共役勾配法

$$\begin{aligned} & \mathbf{x}_0 = \mathbf{0}, \mathbf{p}_0 = \mathbf{b}, \mathbf{r}_0 = \mathbf{b} \\ & \text{for } n = 0, 1, \dots \\ & \quad \alpha_n = \frac{(\mathbf{r}_n, \mathbf{p}_n)}{(\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \text{ or } \frac{(\mathbf{r}_n, \mathbf{r}_n)}{(\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \\ & \quad \mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n \\ & \quad \mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n \mathbf{A}\mathbf{p}_n \\ & \quad \beta_n = -\frac{(\mathbf{r}_{n+1}, \mathbf{A}\mathbf{p}_n)}{(\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \text{ or } \frac{(\mathbf{r}_{n+1}, \mathbf{r}_{n+1})}{(\mathbf{r}_n, \mathbf{r}_n)} \\ & \quad \mathbf{p}_{n+1} = \mathbf{r}_{n+1} + \beta_n \mathbf{p}_n \\ & \text{end} \end{aligned}$$

アルゴリズム 2:  
前処理付き共役勾配法

$$\begin{aligned} & \mathbf{x}_0 = \mathbf{0}, \mathbf{p}_0 = \mathbf{b}, \mathbf{r}_0 = \mathbf{b} \\ & \text{for } n = 0, 1, \dots \\ & \quad \alpha_n = \frac{(\mathbf{r}_n, \mathbf{p}_n)}{(\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \\ & \quad \mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n \\ & \quad \mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n \mathbf{A}\mathbf{p}_n \\ & \quad \widetilde{\mathbf{r}}_{n+1} = \mathbf{K}\mathbf{r}_{n+1} \\ & \quad \beta_n = -\frac{(\widetilde{\mathbf{r}}_{n+1}, \mathbf{A}\mathbf{p}_n)}{(\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \\ & \quad \mathbf{p}_{n+1} = \widetilde{\mathbf{r}}_{n+1} + \beta_n \mathbf{p}_n \\ & \text{end} \end{aligned}$$

線形方程式系 (3) に対し共役勾配法を適用することを前処理付き共役勾配法という。この前処理付き共役勾配法を、さらに数学的に同値なアルゴリズムに変形したものをアルゴリズム 2 で示した (変形については [5] などを参照)。なお、アルゴリズム 2 中の  $\mathbf{K}$  は  $\mathbf{K} = \mathbf{M}^* \mathbf{M}$  で定義される行列である。ただし、この行列  $\mathbf{K}$  を実際にメモリ上に保持する必要はなく、 $\mathbf{r}_{n+1}$  に対する線形変換  $\mathbf{K}\mathbf{r}_{n+1}$  が計算できればよい ( $\mathbf{K}$  の正定値性は必要である)。前処理付き共役勾配法を用いた場合、1 反復の計算量は通常の共役勾配法に比べ、線形変換  $\mathbf{K}\mathbf{r}_{n+1}$  を行う分だけ増える。このため、前処理付き共役勾配法の 1 反復当たりの計算量を増やさないためには、線形変換  $\mathbf{K}\mathbf{r}_{n+1}$  を少ない計算時間とメモリ量で行う必要がある。また、前節で述べたように行列  $\mathbf{M}\mathbf{A}\mathbf{M}^*$  の固有値がより密集しているほうが、前処理付き共役勾配法は高速に収束するため、 $\mathbf{K}$  が  $\mathbf{A}^{-1}$  の近似行列であることが望ましい。これらの性質をふまえた上で、具体的な前処理法については次章で述べる。

## 2.2 共役残差法

共役残差法は、係数行列が正定値行列である線形方程式系を解くクリロフ部分空間法である。この方法の  $n$  回目の反復点  $\mathbf{x}_n$  は、

$$\min_{\mathbf{x} \in \mathcal{K}_n(\mathbf{A}; \mathbf{b})} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$$

の解となっていることが知られている。共役勾配法の場合と同様に、この性質は理論的には共役残差法の  $m$  回での収束性を保証するものの、実際には誤差の影響を受けるため理論通りにはならない。アルゴリズム 3 に共役残差法を示した。

アルゴリズム 3:  
共役残差法

$x_0 = 0, p_0 = b, r_0 = b$   
for  $n = 0, 1, \dots$   
 $\alpha_n = \frac{(r_n, Ap_n)}{(Ap_n, Ap_n)}$   
 $x_{n+1} = x_n + \alpha_n p_n$   
 $r_{n+1} = r_n - \alpha_n Ap_n$   
 $\beta_n = -\frac{(Ar_{n+1}, Ap_n)}{(Ap_n, Ap_n)}$   
 $p_{n+1} = r_{n+1} + \beta_n p_n$   
end

アルゴリズム 4:  
前処理付き共役残差法

$x_0 = 0, p_0 = b, r_0 = b$   
for  $n = 0, 1, \dots$   
 $\alpha_n = \frac{(\widetilde{r}_n, Ap_n)}{(Ap_n, \widetilde{Ap}_n)}$   
 $x_{n+1} = x_n + \alpha_n p_n$   
 $r_{n+1} = r_n - \alpha_n Ap_n$   
 $\widetilde{r}_{n+1} = Kr_{n+1}$   
 $\beta_n = -\frac{(Ar_{n+1}, \widetilde{Ap}_n)}{(Ap_n, \widetilde{Ap}_n)}$   
 $\widetilde{p}_{n+1} = \widetilde{r}_{n+1} + \beta_n p_n$   
 $\widetilde{Ap}_{n+1} = KAp_{n+1}$   
end

アルゴリズム 5:  
前処理付き共役残差法 (実用版)

$x_0 = 0, p_0 = b, r_0 = b$   
for  $n = 0, 1, \dots$   
 $\alpha_n = \frac{(\widetilde{z}_n, q_n)}{(q_n, \widetilde{q}_n)}$   
 $x_{n+1} = x_n + \alpha_n p_n$   
 $r_{n+1} = r_n - \alpha_n q_n$   
 $\widetilde{r}_{n+1} = \widetilde{r}_n - \alpha_n \widetilde{q}_n$   
 $z_{n+1} = Ar_{n+1}$   
 $\beta_n = -\frac{(z_{n+1}, \widetilde{q}_n)}{(q_n, \widetilde{q}_n)}$   
 $p_{n+1} = \widetilde{r}_{n+1} + \beta_n p_n$   
 $q_{n+1} = z_{n+1} + \beta_n q_n$   
 $\widetilde{q}_{n+1} = Kq_{n+1}$   
end

線形方程式系 (3) に対し共役残差法を適用することを前処理付き共役残差法という。この方法をアルゴリズム 4 で示した。なお、アルゴリズム 4 中の  $K$  は  $K = M^*M$  で定義する。アルゴリズム 4 より、共役残差法の各反復では、係数行列  $A$  の線形変換を 2 回と、前処理行列  $K$  の線形変換を 2 回行うことになる。このため、アルゴリズム 4 の計算は前処理付き共役勾配法の計算の 2 倍かかることになる。しかし、[12] で述べたように、

$$\begin{aligned}\widetilde{r}_{n+1} &= Kr_{n+1} \\ &= K(r_n - \alpha_n Ap_n) \\ &= \widetilde{r}_n - \alpha_n KAp_n\end{aligned}$$

という関係を利用することにより、 $\widetilde{r}_{n+1}$  は  $\widetilde{r}_n$  と  $KAp_n$  から更新することができる。また、

$$\begin{aligned}Ap_{n+1} &= A(\widetilde{r}_{n+1} + \beta_n p_n) \\ &= Ar_{n+1} + \beta_n Ap_n\end{aligned}$$

という関係を利用することにより、 $Ap_{n+1}$  は  $Ar_{n+1}$  と  $p_n$  から更新することが可能である。この更新法で計算することにより、前処理付き共役残差法の 1 反復当たりの計算量は、共役勾配法の 1 反復当たりの計算量と同程度にすることができる。この方法を実用版の前処理付き共役残差法としてアルゴリズム 5 に示した。共役勾配法に対する前処理の場合と同様に、行列  $K$  は正定値行列であることが必要であるが、アルゴリズムの実行において実際にメモリ上に保持する必要はなく、線形変換  $Kq_{n+1}$  が行えれば良い。このとき、なるべく少ない計算量とメモリ量で計算できることが必要である。さらに、共役残差法の収束性を高めるためには、 $K$  が  $A^{-1}$  の近似行列であることが望ましい。

### 2.3 双共役勾配法

本節では、双共役勾配法について述べる。共役勾配法が係数行列が正定値である線形方程式系に対する解法であったのに対し、この方法は正定値性を仮定しない一般の係数行列を持つ線形方程式系に対する解法である。次に述べるアルゴリズムから想像出来るように、これは共役勾配法を拡張した方法である。また、この双共役勾配法から派生した方法として2乗共役勾配法、安定化共役勾配法、疑似最小残差法などがあるが、詳細については[2]などを参照されたい。アルゴリズム6で、線形方程式系(1)を解く双共役勾配法を示した。

アルゴリズム6:  
双共役勾配法

```

 $x_0 = 0, p_0 = b, r_0 = b$ 
 $p'_0 = b, r'_0 = b$ 
for  $n = 0, 1, \dots$ 
   $\alpha_n = \frac{(r'_n, r_n)}{(p'_n, Ap_n)}$ 
   $x_{n+1} = x_n + \alpha_n p_n$ 
   $r_{n+1} = r_n - \alpha_n Ap_n$ 
   $r'_{n+1} = r'_n - \bar{\alpha}_n A^* p'_n$ 
   $\beta_n = \frac{(r'_{n+1}, r_{n+1})}{(r'_n, r_n)}$ 
   $p_{n+1} = r_{n+1} + \beta_n p_n$ 
   $p'_{n+1} = r'_{n+1} + \bar{\beta}_n p'_n$ 
end

```

アルゴリズム7:  
前処理付き双共役勾配法

```

 $x_0 = 0, p_0 = b, r_0 = b$ 
 $p'_0 = b, r'_0 = b$ 
for  $n = 0, 1, \dots$ 
   $\alpha_n = \frac{(r'_n, \widetilde{r}_n)}{(p'_n, Ap_n)}$ 
   $x_{n+1} = x_n + \alpha_n p_n$ 
   $r_{n+1} = r_n - \alpha_n Ap_n$ 
   $r'_{n+1} = r'_n - \bar{\alpha}_n A^* p'_n$ 
   $\widetilde{r}_{n+1} = Kr_{n+1}$ 
   $r'_{n+1} = K^* r'_{n+1}$ 
   $\beta_n = \frac{(r'_{n+1}, \widetilde{r}_{n+1})}{(r'_n, \widetilde{r}_n)}$ 
   $p_{n+1} = \widetilde{r}_{n+1} + \beta_n p_n$ 
   $p'_{n+1} = r'_{n+1} + \bar{\beta}_n p'_n$ 
end

```

前処理を施した線形方程式系(2)に対する双共役勾配法と数学的に同値なアルゴリズムがアルゴリズム7である。ここで、アルゴリズム7中の $K$ は、 $K = M_2 M_1$ で定義された行列である。アルゴリズム7では、 $K$ のみが定まれば、前処理付き双共役勾配法のアルゴリズムは前処理行列 $M_1, M_2$ には依存しない。また、これまでと同様に、行列 $K$ は実際にメモリ上に保持する必要はなく、あるベクトル $v$ に対する線形変換 $Kv$ が行えれば、アルゴリズム7を実行することができる。このとき、計算効率の観点から、なるべく少ない計算量とメモリ量で計算できることが必要である。さらに、前処理付き双共役勾配法の収束性を高めるためには、 $K$ が $A^{-1}$ の近似行列であることが望ましい。

### 2.4 一般化共役残差法

本節では、一般化共役残差法について説明する。一般化共役残差法は共役残差法を拡張した方法であり、正定値とは限らない一般の行列を係数行列を持つ線形方程式系を解く。

アルゴリズムの形状は異なるものの、この一般化共役残差法と一般化最小残差法は等しい反復点  $x_n$  を生成することが知られている。また、一般化共役残差法や一般化最小残差法から派生した方法として、途中でリスタートを行うリスタート版一般化共役残差法やリスタート版一般化最小残差法があるが、それらの詳細については [2] などを参照されたい。一般化共役残差法を、アルゴリズム 8 で示した。

アルゴリズム 8 :  
一般共役残差法

$$\begin{aligned} & \mathbf{x}_0 = \mathbf{0}, \mathbf{p}_0 = \mathbf{b}, \mathbf{r}_0 = \mathbf{b} \\ & \text{for } n = 0, 1, \dots \\ & \quad \alpha_n = \frac{(\mathbf{r}_n, \mathbf{A}\mathbf{p}_n)}{(\mathbf{A}\mathbf{p}_n, \mathbf{A}\mathbf{p}_n)} \\ & \quad \mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n \\ & \quad \mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n \mathbf{A}\mathbf{p}_n \\ & \quad \beta_{kn} = -\frac{(\mathbf{A}\mathbf{r}_{n+1}, \mathbf{A}\mathbf{p}_k)}{(\mathbf{A}\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} \\ & \quad \mathbf{p}_{n+1} = \mathbf{r}_{n+1} + \sum_{k=1}^n \beta_{kn} \mathbf{p}_k \\ & \text{end} \end{aligned}$$

アルゴリズム 9 :  
前処理付き一般共役残差法

$$\begin{aligned} & \mathbf{x}_0 = \mathbf{0}, \mathbf{p}_0 = \mathbf{b}, \mathbf{r}_0 = \mathbf{b} \\ & \text{for } n = 0, 1, \dots \\ & \quad \alpha_n = \frac{(\mathbf{r}_n, \mathbf{M}_1^* \mathbf{M}_1 \mathbf{A}\mathbf{p}_n)}{(\mathbf{A}\mathbf{p}_n, \mathbf{M}_1^* \mathbf{M}_1 \mathbf{A}\mathbf{p}_n)} \\ & \quad \mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n \\ & \quad \mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n \mathbf{A}\mathbf{p}_n \\ & \quad \widetilde{\mathbf{r}}_{n+1} = \mathbf{K} \mathbf{r}_{n+1} \\ & \quad \beta_{kn} = -\frac{(\mathbf{A}\widetilde{\mathbf{r}}_{n+1}, \mathbf{M}_1^* \mathbf{M}_1 \mathbf{A}\mathbf{p}_k)}{(\mathbf{A}\mathbf{p}_k, \mathbf{M}_1^* \mathbf{M}_1 \mathbf{A}\mathbf{p}_k)} \\ & \quad \mathbf{p}_{n+1} = \widetilde{\mathbf{r}}_{n+1} + \sum_{k=1}^n \beta_{kn} \mathbf{p}_k \\ & \text{end} \end{aligned}$$

これまでと同様に、前処理を施した線形方程式系 (2) に対し一般共役残差法を適用するものと数学的に同値となる解法をアルゴリズム 9 で示す。アルゴリズム 9 中の  $\mathbf{K}$  は  $\mathbf{K} = \mathbf{M}_2 \mathbf{M}_1$  で定義される行列である。アルゴリズム 7 の前処理付き双共役勾配法の場合と違い、線形変換  $\mathbf{K}\mathbf{v}$  の他に線形変換  $\mathbf{M}_1^* \mathbf{M}_1 \mathbf{v}$  の計算が必要であることに注意する。つまり、前処理付き一般化共役残差法では、 $\mathbf{K}$  だけでなく  $\mathbf{M}_1, \mathbf{M}_2$  に依存して、アルゴリズムは変化する。

### 3 様々な前処理行列

本章では、幾つかの前処理について述べる。前章で述べたように、前処理行列  $\mathbf{K} = \mathbf{M}_2 \mathbf{M}_1$  に求められている性質として、なるべく少ない計算量とメモリ量で線形変換  $\mathbf{K}\mathbf{v}$  が計算できることと、 $\mathbf{K}$  が  $\mathbf{A}^{-1}$  の近似行列であることが挙げられる。このことを踏まえた上で、以下の節では対角スケールリング、不完全コレスキー分解、多項式前処理について説明する。一般的に、良い前処理とは係数行列の性質や使用できるメモリの量などに強く依存するため、決定的な前処理法は存在しない。

### 3.1 対角スケーリング

一番簡単な前処理として、対角スケーリングがある。これは、行列  $A$  の対角要素を抜き出した行列を  $D$  とすると、前処理行列  $K$  として  $D^{-1}$  とすることに相当する。 $D$  が  $A$  の近似行列であるとみなせば、 $K$  は  $A^{-1}$  の近似行列といえる。一般的にこの前処理行列  $K$  は  $A^{-1}$  のそれほど良い近似ではないため、収束性に関して大きな効果が期待できない。しかし、線形変換  $Kv$  には  $m$  flops しかかからず、計算量が少ないため良く利用される。さらに、 $A$  が正定値なら  $K$  も正定値となる性質があるため、共役勾配法や共役残差法の前処理として利用できる。

### 3.2 不完全コレスキー分解

係数行列  $A$  が正定値である場合によく使われる前処理法として不完全コレスキー分解がある。これは、係数行列  $A$  のゼロの部分が分解した下三角行列  $L$  でもゼロとなるように、近似的にコレスキー分解を行う方法である。不完全コレスキー分解は以下のようにして計算できる。

```

for k = 1, 2, ..., n
  Lkk = √Lkk
  for i = k + 1, k + 2, ..., n    Aik ≠ 0 の場合
    Lik = Lik/Lkk
  end
  for j = k + 1, k + 2, ..., n
    for i = j, j + 1, ..., n    Aij ≠ 0 の場合
      Lij = Lij - LikLjk
    end
  end
end
end

```

この不完全コレスキー分解  $A \approx LL^*$  を利用し、前処理行列として  $K = L^{-*}L^{-1}$  と定める。このとき、 $K \approx A^{-1}$  となるため、 $K$  は  $A^{-1}$  の近似行列とみなせる。下三角行列  $L$  の構成法により、 $A$  と  $L$  は同程度に疎である。しかし、 $L^{-*}L^{-1}(=K)$  は疎であるとは限らない。このため実用上は、疎な三角行列  $L$  を保持した上で、線形変換  $Kv$  を線形方程式系  $LL^*z = v$  を解くことにより求める。この線形方程式系の解は、 $L$  が三角行列であるため  $L$  の疎性を直接利用して計算できる。この場合、 $A$  と  $L$  の疎性が等しいことを考慮すると、線形変換  $Av$  と同じくらいの計算量で前処理が可能である。また、 $L$  が生成できれば  $K$  は正定値行列となるため、共役勾配法や共役残差法の前処理として利用できる。しかし、 $A$  が正定値でも不完全コレスキー分解ができるとは限らない。その場合、 $A + \epsilon I$  を不完全コレスキー分解する方法などが提案されている。また、係数行列が正定値ではない場合、不完全コレスキー分解の代わりに不完全 LU 分解が使われることが多い [3]。

### 3.3 多項式近似

多項式近似とは、前処理行列  $K$  は  $A$  の多項式で近似する方法である。一般に、 $M$  を正則行列としたとき、 $G = I - M^{-1}A$  と定義すると、 $\rho(G) < 1$  ならば、

$$A^{-1} = M^{-1} + GM^{-1} + G^2M^{-1} + G^3M^{-1} + \dots$$

が成り立つことが知られている [5]。これは無限級数展開であるが、これを  $t$  まで展開したものを前処理行列として利用することを考える。そのような行列を  $K_t$  とすると、つまり

$$K_t = M^{-1} + GM^{-1} + G^2M^{-1} + \dots + G^{t-1}M^{-1}$$

と書くことができる。クリロフ部分空間法では、あるベクトル  $v$  に対する線形変換  $K_tv$  ができればよく、 $K_t$  は陽に行列を生成する必要はない。ここで、

$$\begin{aligned} K_tv &= (GK_{t-1} + M^{-1})v \\ &= M^{-1}(r - AK_{t-1}v) + K_{t-1}v \end{aligned}$$

という関係が成り立つため、

$$z_0 = v, \quad z_k = M^{-1}(r - Az_{k-1}) + z_{k-1} \quad (5)$$

という漸化式を添字が  $t$  まで行うことにより、線形変換  $K_tv$  を求めることができる。つまり、(5) は定常反復法と見なすことが可能である。この計算を効率良く行うためには、行列  $M^{-1}$  とベクトルの積を少ない計算量で行う必要がある。ここで、 $A$  の対角部分を抜き出した行列を  $D$ 、狭義下三角部分を抜き出した行列を  $L$ 、狭義上三角部分を抜き出した行列を  $U$  とする。すると、 $M$  として  $D$  を取るとこれはヤコビ法となる。また、 $D + L$  あるいは  $D + U$  を取ると、これはガウス・ザイデル法となる。さらに、 $\frac{1}{\omega}D + L$  あるいは  $\frac{1}{\omega}D + U$  を取ると、これは逐次過剰緩和法 (SOR 法) となる。このような特別な場合には、より効率よく線形変換  $K_tv$  を計算することができる。さらに [1] では、一般化共役残差法に対する可変の前処理という枠組みを提案している。その中では、前処理として SOR 法を用いた数値実験が行われている。SOR 法などを前処理として利用した場合、たとえ係数行列  $A$  が正定値の場合でも、前処理行列  $K$  の正定値性は失われる。この問題を解決するため、対称逐次過剰緩和法 (SSOR 法) などの方法が提案されている [5]。

## 4 係数行列が密の場合

クリロフ部分空間法は係数行列の疎性を崩さない方法であるため、特に係数行列が疎である線形方程式系に対し利用されることが多い。一方、係数行列が特別な構造を持っている場合、その構造を利用することにより係数行列が密である線形方程式系に対し効率良く働く可能性がある。本章では、半正定値計画問題を主双対内点法で解く場合、その途中で現れる線形方程式系に対しクリロフ部分空間法が有効であることを述べる。



#### 4.1 半正定値計画問題とは

まず、 $S^n$  を  $n \times n$  の実対称行列の集合と定義する。また、 $S_+^n$  を  $n \times n$  の半正定値対称行列の集合と定義する。任意の  $X, Z \in \mathfrak{R}^{n \times n}$  に対して、 $X \bullet Z$  は  $X$  と  $Z$  の内積、すなわち、 $\text{Tr } X^T Z$  ( $X^T Z$  のトレース) を表す。

$F_i \in S^n$  ( $i = 0, 1, \dots, m$ ),  $f_i \in \mathfrak{R}$  ( $i = 1, 2, \dots, m$ ) とする。このとき、半正定値計画 (Semidefinite Programming: SDP) の主問題と双対問題は以下のように与えられる。

$$\left. \begin{array}{l} \text{主問題:} \\ \text{最小化} \quad F_0 \bullet X \\ \text{制約条件} \quad F_i \bullet X = f_i \quad (i = 1, 2, \dots, m), \\ \quad \quad \quad X \in S_+^n. \end{array} \right\} \quad (6)$$

$$\left. \begin{array}{l} \text{双対問題:} \\ \text{最大化} \quad \sum_{i=1}^m f_i y_i \\ \text{制約条件} \quad \sum_{i=1}^m F_i y_i + Z = F_0, \\ \quad \quad \quad Z \in S_+^n. \end{array} \right\} \quad (7)$$

この問題は線形計画問題の対称行列の空間への拡張として捉えることができ、線形計画問題の主双対内点法を拡張した内点法で理論的に多項式時間内で解くことが可能である [7, 10, 11]。この主双対内点法では、探索方向の計算過程である種の Newton 法を使用しており、このとき、サイズが  $m$  (主問題の線形制約の数) の線形方程式系

$$Acy = b \quad (8)$$

を解く必要が生ずる。ここで、線形方程式系 (8) の係数行列  $A$  の各要素は、

$$A_{ij} = F_i \bullet X F_j Z^{-1} \quad (i, j = 1, 2, \dots, m),$$

と定義される。この係数行列  $A$  に関し、重要な 2 つの性質が知られている [8]。一つは、係数行列  $A$  が正定値行列になっていることであり、もう一つは、一般に行列  $X, Z$  が密であるため  $F_i$  ( $i = 1, 2, \dots, m$ ) が疎行列であっても、係数行列  $A$  は密行列となることである。このため、主問題の線形制約の数  $m$  が非常に多い SDP 問題を解く場合、非常に大きな次元の係数行列が密な線形方程式系を解く必要が生ずる。このため、直接法では計算量やメモリ使用量の制約からこの大規模線形方程式系の解の計算が困難となる。そこで、主双対内点法の各反復で解く線形方程式系 (8) に共役勾配法を適用することを試みる。

## 4.2 係数行列とベクトルの計算

共役勾配法では、係数行列  $A$  はベクトル  $p$  に対する線形変換にのみ利用する。この線形変換  $Ap$  となるベクトルを  $q$  とすると、任意の  $i$  に対して、

$$\begin{aligned} q_i &= \sum_{j=1}^m A_{ij} p_j = \sum_{j=1}^m \mathbf{F}_i \bullet \mathbf{X} \mathbf{F}_j \mathbf{Z}^{-1} p_j \\ &= \sum_{j=1}^m \text{Tr}(\mathbf{F}_i \mathbf{X} \mathbf{F}_j \mathbf{Z}^{-1}) p_j \\ &= \text{Tr}(\mathbf{F}_i \mathbf{X} \sum_{j=1}^m \mathbf{F}_j p_j \mathbf{Z}^{-1}) \\ &= \mathbf{F}_i \bullet \mathbf{X} \sum_{j=1}^m \mathbf{F}_j p_j \mathbf{Z}^{-1} \end{aligned}$$

が成り立つ。よって、以下のような手順で線形変換  $Ap$  の計算を行うことができる。

$$\mathbf{H}_1 = \sum_{j=1}^m p_j \mathbf{F}_j, \quad \mathbf{H}_2 = \mathbf{X} \mathbf{H}_1 \mathbf{Z}^{-1}, \quad q_i = \mathbf{F}_i \bullet \mathbf{H}_2 \quad (i = 1, 2, \dots, m) \quad (9)$$

この計算方法では、計算量としてほぼ大きさ  $n \times n$  の行列同士のかけ算 2 回分となる。 $m = \mathcal{O}(n^2)$  であることを考慮すると、この計算は  $\mathcal{O}(m^{1.5})$  flops と見積もることができる。係数行列  $A$  は  $m^2$  個の (ゼロでない) 要素を持つのに対し、係数行列  $A$  の構造を利用することにより線形変換を  $\mathcal{O}(m^{1.5})$  flops で計算することが可能となった。また、 $m \times m$  の密行列  $A$  を保持せず線形変換  $Ap$  を計算している。

## 4.3 多項式近似を利用した前処理

次に、共役勾配法に対する前処理について考察する。線形方程式 (8) の係数行列  $A$  は密であるため、前処理として不完全コレスキー分解を利用することはできない。前節では  $A$  の構造を利用することにより線形変換  $Ap$  を効率よく計算する方法を説明したが、前処理においても  $A$  の構造を利用したい。このため、3.3 節で説明した多項式近似を利用することが適切であると思われる。前処理としてヤコビ法を利用する場合、前節で述べた計算法を直接利用することにより、高速に前処理を行うことができる。しかし、前処理として SOR 法や SSOR 法を利用する場合、前節と同様の方法では計算できない。このため以下の工夫を行う。

SOR 法の  $k$  反復目の反復点  $x$  の  $i$  番目の要素を  $x_i^{(k)}$  とする。このとき、

$$\begin{aligned} x_i^{(k+1)} &= \sum_{j=1}^{i-1} A_{ij} x_j^{(k+1)} + \sum_{j=i+1}^m A_{ij} x_j^{(k)} \\ &= \sum_{j=1}^{i-1} (\mathbf{F}_i \mathbf{Z}^{-1} \bullet \mathbf{X} \mathbf{F}_j) x_j^{(k+1)} + \sum_{j=i+1}^m (\mathbf{F}_i \mathbf{Z}^{-1} \bullet \mathbf{X} \mathbf{F}_j) x_j^{(k)} \\ &= \mathbf{F}_i \mathbf{Z}^{-1} \bullet (\mathbf{X} (\sum_{j=1}^{i-1} x_j^{(k+1)} \mathbf{F}_j) + \mathbf{X} (\sum_{j=i+1}^m x_j^{(k)} \mathbf{F}_j)) \end{aligned}$$

が成り立つ。ここで、 $H_i = X(\sum_{j=1}^{i-1} x_j^{(k+1)} F_j) + X(\sum_{j=i+1}^m x_j^{(k)} F_j)$  と定義すると、この  $H_i$  は  $H_i = H_{i-1} + x_{i-1}^{(k+1)} X F_{i-1} - x_i^{(k)} X F_i$  という漸化式で書き表せる。この事実から、以下のような手順により SOR 法を実行することができる。

```

x = 0, H = O
for k = 0, 1, ...
  for i = 1, 2, ..., m
    H = H + x_{i-1}^{(k+1)} X F_{i-1} - x_i^{(k)} X F_i
    (もし i = 1 のとき H = H + x_m^{(k)} X F_m - x_1^{(k)} X F_1)
    J = F_i Z^{-1}
    x_i^{(k+1)} = ω(b_i - J • H) / A_{ii} + (1 - ω)x_i^{(k)}
  end
end

```

このアルゴリズムの詳細については [9] を参照されたい。結果として、SOR 法の 1 反復を  $O(n^3)$  で計算することができる。これは、 $A_p$  の計算にかかる計算量と同等であり、実用的な前処理となる。

## 謝辞

本研究をするにあたり、東京大学工学系研究科の張紹良助教授と新田義寛さんの助力を得ました。ここに、感謝の意を表します。

## 参考文献

- [1] 阿部邦美, 張紹良, 長谷川秀彦, 姫野龍太郎, SOR 法を用いた可変的前処理付き一般化共役残差法, 応用数理学会 論文誌, 第 11 巻, 4 号 (2001) 157-170.
- [2] R. Barrett, M. Barry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Römme, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods* (Society for Industrial and Applied Mathematics, 1994)
- [3] 藤野 清次, 張 紹良, 反復法の数理, (朝倉書店, 1996)
- [4] K. Fujisawa, M. Fukuda, M. Kojima and K. Nakata, Numerical evaluation of SDPA (SemiDefinite Programming Algorithm), in: H. Frenk, K. Roos, T. Terlaky and S. Zhang, eds., *High Performance Optimization* (Kluwer Academic Publishers, Dordrecht, 1999) 267-301.

- [5] G. H. Golub and C. F. Van Loan, *Matrix Computations* (The John Hopkins University Press, Baltimore, Maryland, 1983).
- [6] M. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, *Journal Research of the National Bureau of Standards*, 49 (1952) 409–436.
- [7] 小島政和, 半正定値計画とその組合せ最適化への応用, 離散構造とアルゴリズム 5, (近代科学社, 1998).
- [8] 中田 和秀, 藤沢 克樹, 小島 政和, 半正定値計画問題に対する主双対内点法における共役勾配法の実装, 統計数理 第46巻 第2号 (1998) 297–316.
- [9] 新田 義寛, 半正定値計画で現れる線形方程式系に対する前処理付き共役残差法, 東京大学 物理工学専攻 修士論文, (2002).
- [10] L. Vandenberghe and S. Boyd, Semidefinite Programming, *SIAM Review* 38 (1996) 49–95.
- [11] H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications* (Kluwer Academic Publishers, Massachusetts, USA, 2000).
- [12] S.-L. Zhang, K. Nakata and M. Kojima, Incomplete orthogonalization preconditioner for solving large and dense linear systems which arise from semidefinite programming, *Applied Numerical Mathematics*, 41 (2002) 235–245.