

Krylov 部分空間法

情報科学類 3 年 江畑 拓哉 (201611350)

Outline

- 次元削減を行うために、切り詰めた (truncated) SVD (特異値分解) や PCR (主成分回帰) を用いることがある。
この時、右辺はどのような基底ベクトル z_i の選択にも影響を与えない。
⇒ [Example 7.2] 残差の減衰率は q_2 の方が q_1 に比べてかなり緩くなっている。
- 多くの場合、任意の右辺の基底ベクトルの数で作られる関数としては、残差は急速に減衰してほしい。
つまり右辺には、基底ベクトルの選択に何らかの影響を与えて欲しい。
- この解法として Lanczos-Golub-Kahan (LGK) 二重対角化 (bidiagonalization) がある。

用語

- 特異値分解

任意の \mathbf{A} where $\mathbf{A} \in \mathbb{R}^{m \times n}$ に対して

$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$ where $\mathbf{U} \in \mathbb{R}^{m \times m}$ is orthogonal

$\mathbf{V} \in \mathbb{R}^{n \times n}$ is orthogonal

$\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ is diagonal $\sigma_1 \geq \dots \geq \sigma_r > 0$

となる $\mathbf{\Sigma}$ を求めること

- Truncated SVD

SVD の近似

- 主成分回帰

独立変数から共分散行列を求め、固有値分解をして主成分を抽出し (PCA(主成分分析))、その主成分を用いて回帰分析をする手法

LGK 二重対角化

- 線形代数の分野で用いられていおり、 Lanczos 二重対角化とも呼ばれる。
- 計量科学やその他の分野では部分的最小二乗法 (PLS(Partial Least Squares/Projection Latent structures)) に密接して関連した方法として知られている。
- Krylov 部分空間法 の中から来た手法であり、疎な線形問題 (sparse linear systems) や、固有値・特異値の計算で用いられる。

Krylov 部分空間法は再帰的であるが、始めに導出として扱うのはハウスホルダー変換を用いて行列を二重対角化する手法である。

ハウスホルダー変換を用いた二重対角化

密行列 $C \in \mathbb{R}^{m \times (n+1)}$ の SVD を計算するアルゴリズムの最初のステップは、左から右へハウスホルダー変換をし、上二重対角の形に C を変換することである。

$m > n$ を仮定すると、以下の形が求まる。

$$C = P \begin{pmatrix} \hat{B} \\ 0 \end{pmatrix} W^T$$

where P, W is orthogonal

B is upperbidiagonal (7.5)

この自己圧縮は他の目的でも有用。

⇒ 疎・密行列の最小二乗問題の近似解を求める際など

ハウスホルダー変換を用いた二重対角化の例 I

例として、 $\mathbf{C} \in \mathbb{R}^{6 \times 5}$ を用いて説明する。※ $m = 6, n + 1 = 5$
まず 1 列目の非対角成分を、左側からかける変換行列 $\mathbf{P}_1^T \in \mathbb{R}^{6 \times 6}$ を用いて 0 にする。

$$\mathbf{P}_1^T \mathbf{C} = \mathbf{P}_1^T \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{pmatrix} = \begin{pmatrix} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{pmatrix}$$

次に、別の右からかける変換行列 \mathbf{W}_1 を用いて 1 行目の 3 番目の要素から右側をすべて 0 にします。 \mathbf{W}_1 は以下のように表すことができる。

$$\mathbb{R}^{5 \times 5} \ni \mathbf{W}_1 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_1 \end{pmatrix} \text{ where } \mathbf{Z}_1 \text{ is Householder transformation}$$

これを用いた変換は 1 列目の要素に対しては何も影響しない。つまり前に行った 1 列目の非対角成分を 0 にした変換の効果は打ち消されない。結果は以下ようになる。

$$\mathbf{P}_1^T \mathbf{C} \mathbf{W}_1 = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{pmatrix} \mathbf{W}_1 = \begin{pmatrix} \times & * & 0 & 0 & 0 \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{pmatrix} =: \mathbf{C}_1$$

同様の手法を用いて 2 列目の対角成分よりも下の成分を 0 にする変換行列 \mathbf{P}_2 を \mathbf{C}_1 にかける。但し 1 行目の要素については値が変わらないようにする。

すると、 \mathbf{P}_2 は以下のように表すことができる。

$$\mathbb{R}^{6 \times 6} \ni \mathbf{P}_2 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{P}}_2 \end{pmatrix} \text{ where } \tilde{\mathbf{P}}_2 \in \mathbb{R}^{5 \times 5} \text{ is Householder transformation}$$

これを適用して、

$$\mathbf{P}_2^T \mathbf{C}_1 = \begin{pmatrix} \times & \times & 0 & 0 & 0 \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix}$$

そして同様に右からかける変換行列 \mathbf{W}_2 は以下のように表すことができる。

$$\mathbf{W}_2 = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \end{pmatrix}, \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

これも \mathbf{W}_1 同様 \mathbf{P}_2 で 0 にした部分を汚すことなく 2 列目の 4 行目以降の要素を 0 にする。

$$\mathbf{P}_2^T \mathbf{C}_1 \mathbf{W}_2 = \begin{pmatrix} \times & \times & 0 & 0 & 0 \\ 0 & \times & * & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix} =: \mathbf{C}_2$$

この操作を繰り返すことで最終的に以下の形を得る。

$$P^T C W = \begin{pmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times & \times \\ & & & & & \times \end{pmatrix} = \begin{pmatrix} \hat{B} \\ 0 \end{pmatrix}$$

where $P = P_1 P_2 \cdots P_n \in \mathbb{R}^{m \times m}$

$$W = W_1 W_2 \cdots W_{n-2} \in \mathbb{R}^{(n+1) \times (n+1)} \quad (7.6)$$

※ W が $n-2$ までであるのは、上二重対角であるから。

一般的に上で示された P , W はハウスホルダー変換の積であり、

$$\hat{\mathbf{B}} = \begin{pmatrix} \beta_1 & \alpha_1 & & & \\ & \beta_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & & & \beta_n & \alpha_n \\ & & & & \beta_{n+1} \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}$$

は上二重対角行列である。

これらは、この章の残りの部分で用いられる直交行列を生成するための構造を持っている。

命題 7.3

二重対角分解 (7.6) 式における \mathbf{P} の列を $\mathbf{p}_i, i = 1, 2, \dots, m$ として表すと以下のように表すことが出来る。

$$\mathbf{p}_1 = \beta_1 \mathbf{c}_1, \mathbf{W} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix} \text{ where } \mathbf{Z} \in \mathbb{R}^{n \times n} \text{ is orthogonal}$$

\mathbf{c}_1 は \mathbf{C} の最初の列である。

証明

$i = 1$ の場合は、 $\mathbf{P}^T \mathbf{c}_1 = \beta_1 \mathbf{e}_1$ であったことから明らかである。これ以降の場合では、 \mathbf{W}_i が以下の構造で表されていたことからわかる。

$$\mathbf{W}_i = \begin{pmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_i \end{pmatrix}$$

where $\mathbf{I}_i \in \mathbb{R}^{i \times i}$ is identity matrices
 \mathbf{Z}_i are orthogonal

ハウスホルダー変換を用いた二重対角化への削減は $4mn^2 - 4n^3/3$ flops かかる。

もし $m \gg n$ ならば、 \mathbf{A} を上三角行列にして、 \mathbf{R} 要素を二重対角化したほうが良い。

※ \mathbf{R} とは LU 分解で言う \mathbf{U} 要素 (LU 分解の別名は LR 分解)

最小二乗問題 I

最小二乗問題 $\min_x \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ を解く場合について考える。

二重対角化において $\mathbf{C} = (\mathbf{b} \ \mathbf{A})$ とした場合、同等の二重対角化最小二乗問題 (bidiagonal least squares problem) を得ることが出来る。式 (7.6) と命題 7.3 より以下を得る。

$$\mathbf{P}^T \mathbf{C} \mathbf{W} = \mathbf{P}^T (\mathbf{b} \ \mathbf{A}) \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix} = (\mathbf{P}^T \mathbf{b} \ \mathbf{P}^T \mathbf{A} \mathbf{Z}) = \begin{pmatrix} \beta_1 & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (7.7)$$

where

$$\mathbf{B} = \begin{pmatrix} \alpha_1 & & & & \\ \beta_2 & \alpha_2 & & & \\ & \ddots & \ddots & & \\ & & \beta_n & \alpha_n & \\ & & & \beta_{n+1} & \end{pmatrix} \in \mathbb{R}^{(n+1) \times n}$$

そして $\mathbf{y} = \mathbf{Z}^T \mathbf{x}$ として残差の 2 ノルムを以下のように書くことができる。

$$\begin{aligned}\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 &= \left\| \begin{pmatrix} \mathbf{b} & \mathbf{A} \end{pmatrix} \begin{pmatrix} 1 \\ -\mathbf{x} \end{pmatrix} \right\|_2 = \left\| \mathbf{P}^T \begin{pmatrix} \mathbf{b} & \mathbf{A} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix} \begin{pmatrix} 1 \\ -\mathbf{y} \end{pmatrix} \right\|_2 \\ &= \left\| \begin{pmatrix} \mathbf{P}^T \mathbf{b} & \mathbf{P}^T \mathbf{A} \mathbf{Z} \end{pmatrix} \begin{pmatrix} 1 \\ -\mathbf{y} \end{pmatrix} \right\|_2 = \|\beta_1 \mathbf{e}_1 - \mathbf{B}\mathbf{y}\|_2 \quad (7.8)\end{aligned}$$

もし平面回転の操作によって \mathbf{B} が上二重対角行列に変換される場合、この二重対角最小二乗問題 $\min_{\mathbf{y}} \|\beta_1 \mathbf{e}_1 - \mathbf{B}\mathbf{y}\|_2$ は $O(n)$ flops で解くことができる。

LGK 二重対角化

ここでは前で述べた二重対角化手法ではない別の手法を扱う。この手法は式 (7.7) の計算を再帰的に解く。これを LGK 二重対角化 という。式 (7.7) の最後の式は以下のように書くことが出来る。

$$\mathbf{P}^T \mathbf{A} = \begin{pmatrix} \mathbf{BZ}^T \\ \mathbf{0} \end{pmatrix} \text{ where } \mathbf{BZ}^T \in \mathbb{R}^{(n+1) \times n}$$

これは更に詳しく書けば以下のようなになる。

$$\begin{aligned} \mathbf{A}^T (\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_{n+1}) &= \mathbf{Z} \mathbf{B}^T \\ &= (\mathbf{z}_1 \quad \mathbf{z}_2 \quad \cdots \quad \mathbf{z}_n) \begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ & \alpha_2 & \beta_3 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \beta_i & \\ & & & & \alpha_i & \\ & & & & & \ddots & \ddots \\ & & & & & & \alpha_n & \beta_{n+1} \end{pmatrix} \end{aligned}$$

両辺の i 列 ($i \geq 2$) を比較すると以下のような式が考えられる。

$$\mathbf{A}^T \mathbf{p}_i = \beta_i \mathbf{z}_{i-1} + \alpha_i \mathbf{z}_i$$

変形して

$$\alpha_i \mathbf{z}_i = \mathbf{A}^T \mathbf{p}_i - \beta_i \mathbf{z}_{i-1} \tag{7.9}$$

同様に i 列について、 $(\mathbf{P}\mathbf{P}^T\mathbf{A}\mathbf{Z} = \mathbf{P}\mathbf{B}\mathbf{Z}^T\mathbf{Z})$

$$\begin{aligned}\mathbf{A}\mathbf{Z} &= \mathbf{A}(\mathbf{z}_1 \quad \mathbf{z}_2 \quad \cdots \quad \mathbf{z}_n) \\ &= \mathbf{P}\mathbf{B} = (\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_{n+1}) \begin{pmatrix} \alpha_1 & & & & & & & \\ \beta_2 & \alpha_2 & & & & & & \\ & \ddots & \ddots & & & & & \\ & & & \alpha_i & & & & \\ & & & \beta_{i+1} & & & & \\ & & & & \ddots & \ddots & & \\ & & & & & \beta_n & \alpha_n & \\ & & & & & & \beta_{n+1} \end{pmatrix}\end{aligned}$$

同様に以下のような式が考えられる。

$$\mathbf{Az}_i = \alpha_i \mathbf{p}_i + \beta_{i+1} \mathbf{p}_{i+1}$$

変形して

$$\beta_{i+1} \mathbf{p}_{i+1} = \mathbf{Az}_i - \alpha_i \mathbf{p}_i \quad (7.10)$$

命題 7.3 より等式 $\beta_1 \mathbf{p}_1 = \mathbf{b}$ を初期値として等式 (7.9), (7.10) を再帰的に解くことが出来る。

LGK 二重対角化

- 1 $\beta_1 \mathbf{p}_1 = \mathbf{b}, \mathbf{z}_0 = 0$
- 2 *for* $i = 1 : n$
 $\alpha_i \mathbf{z}_i = \mathbf{A}^T \mathbf{p}_i - \beta_i \mathbf{z}_{i-1}$
 $\beta_{i+1} \mathbf{p}_{i+1} = \mathbf{A} \mathbf{z}_i - \alpha_i \mathbf{p}_i$
- 3 *end*

係数 α_{i-1}, β_i は $\|\mathbf{p}_i\| = \|\mathbf{z}_i\| = 1$ となるように決定する。

再帰の終点は α_i 又は β_i がゼロと等しくなったときである。最小二乗問題の解法においてこれは明確に定義されている特別な場合であることがわかっているので、終点とみなしても問題がない。

正確な算術手続きを考えれば、再帰的二重対角化の手続きはハウスホルダー変換を用いた二重対角化と等しくなる。従ってこの過程で生成された $(\mathbf{p}_i)_{i=1}^n$ $(\mathbf{z}_i)_{i=1}^n$ は $\mathbf{p}_i^T \mathbf{p}_j = 0$, $\mathbf{z}_i^T \mathbf{z}_j = 0$ if $i \neq j$ を満たす。

しかし浮動点計算を考慮すると、生成されるベクトルは再帰の過程で直交性を失ってしまう。