数理アルゴリズムとシミュレーション 演習課題 6

提出期限: 2020/11/26 23:59

以下の課題を行い，レポートを提出すること．レポートの作成に関しては，manaba の「演習課題」ページ内の項目をしっかりと確認すること．また，レポートの作成にあたって参考にした文献や Web ページはその出典を明示すること．

課題 1

行列 $A \in \mathbb{R}^{m\times n}$ の特異値分解を

$$A = U\Sigma V^{\mathrm{T}},\ U^{\mathrm{T}}U = UU^{\mathrm{T}} = I, V^{\mathrm{T}}V = VV^{\mathrm{T}} = I$$

とし，$U \in \mathbb{R}^{m\times m}$ の列ベクトルを左特異ベクトル，$V \in \mathbb{R}^{n\times n}$ の列ベクトルを右特異ベクトル，$\Sigma \in \mathbb{R}^{m\times n}$ の対角要素を特異値とする．このとき，以下の問いに答えよ．

(1-1) AA^{T} の固有ベクトルに等しいものを選べ．
(a) 左特異ベクトル　(b) 右特異ベクトル

(1-2) $A^{\mathrm{T}}A$ の固有ベクトルに等しいものを選べ．
(a) 左特異ベクトル　(b) 右特異ベクトル

(1-3) 行列 $A \in \mathbb{R}^{n\times n}$ の特異値を $\sigma_i,\ i = 1, 2, \ldots, n$ とする．このとき, $2n$ 次行列

$$\begin{bmatrix} O & A^H \\ A & O \end{bmatrix}$$

の固有値として正しいものを選べ．
(a) $\sigma_1, \sigma_1^2, \sigma_2, \sigma_2^2, \ldots, \sigma_n, \sigma_n^2$　(b) $-\sigma_1, -\sigma_1^2, -\sigma_2, -\sigma_2^2, \ldots, -\sigma_n, -\sigma_n^2$
(c) $\pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_n$　(d) $\pm\sigma_1^2, \pm\sigma_2^2, \ldots, \pm\sigma_n^2$

課題 2

行列 $A = \{a_{ij}\} \in \mathbb{R}^{100\times 10}$ を以下のように定義する．

$$a_{ij} = |i - j + 1|.$$

このとき，べき乗法，逆反復法を $A^{\mathrm{T}}A$ に適用し，A の最大特異値，最小特異値を求めよ．ただし，初期ベクトルの要素は全て 1 とし，k 反復目で求めた $A^{\mathrm{T}}A$ の近似固有値 $\lambda^{(k)}$ と

それに対応する近似固有ベクトル $\boldsymbol{x}^{(k)}$ による残差

$$r = \left\| A^{\mathrm{T}} A \boldsymbol{x}^{(k)} - \lambda^{(k)} \boldsymbol{x}^{(k)} \right\|_2$$

が 10^{-5} 以下となったときに反復を停止せよ．

課題 3

図 1 のようなデータが与えられる．このデータに対して主成分分析を行い，主成分方向を示した直線を加えてグラフに描画せよ．なお，データは manaba にアップロードしているのでダウンロードして用いること．データの読み込みは関数 **`load`** を用いると行える．データは変数 **`X`**, **`Y`** で構成されており，それぞれデータの x 座標，y 座標を表す．

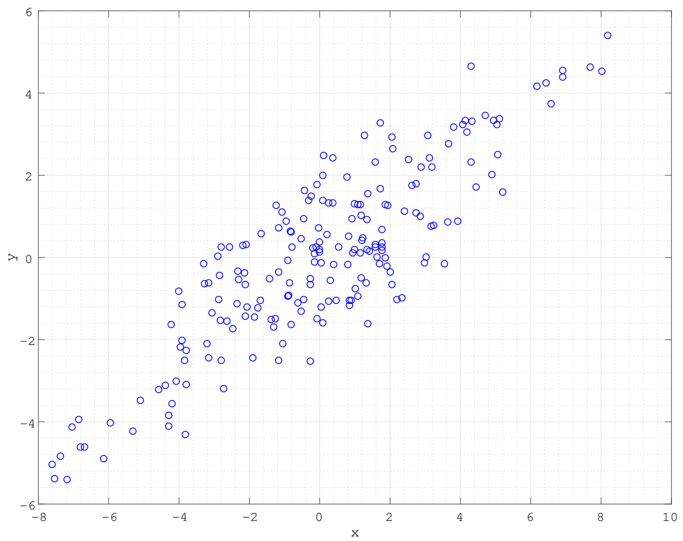

図 1　データの分布

(次ページへ続く)

課題 4

特異値分解を用いると，行列 $A \in \mathbb{R}^{m\times n}$, $m > n$ は以下のように表せる．

$$A = \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}.$$

ここで，A の特異値，左特異ベクトル，右特異ベクトルをそれぞれ $\sigma_i \in \mathbb{R}, \boldsymbol{u}_i \in \mathbb{R}^m, \boldsymbol{v}_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, n$ とする．σ_i の値が小さくなると，A に与える影響が少ないと考えることができるので，ある $k \leq n$ 個の特異値，右特異ベクトル，左特異ベクトルで近似することができる．このとき，近似した A は以下のように表せる．

$$\tilde{A}_k = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}},\ k = 1, 2, \ldots, n.$$

これを，A の低ランク近似と呼ぶ．これを用いて，以下の問いに答えよ．

(4-1) 以下の MATLAB プログラムで行列 A を生成する．

ソースコード 1　行列 A を生成するプログラム

```
rng(0);
U1 = rand(100,10);
V1 = rand(70, 10);

U2 = rand(100, 40);
V2 = rand(70, 40);

A = U1*V1' + 1.0E-5*U2*V2' + 1.0E-10*rand(100,70);
```

このとき，低ランク近似を $k = 1, 2, \ldots, n$ に対して行い，それぞれの k でのフロベニウスノルムに関する誤差 $\|A - \tilde{A}_k\|_{\mathrm{F}}$ を計算し，グラフに描画せよ．なお，$A = \{a_{ij}\} \in \mathbb{R}^{m\times n}$ のフロベニウスノルムは以下のように定義する．

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

A のフロベニウスノルムは, MATLAB の関数 **norm(A,'fro')** を用いて計算できる．また，特異値・特異ベクトルは MATLAB の関数 **svd** を用いて計算してもよい．

(4-2) (4-1) で計算した，それぞれの k に対するフロベニウスノルムに関する誤差 $\|A - \tilde{A}_k\|_{\mathrm{F}}$ は，

$$\sqrt{\sum_{l=k+1}^{n} \sigma_l^2}$$

に一致する．これを，(4-1) で用いた行列 A で確認し，上記の値をグラフに描画せよ．