数値計算 参考問題（本番より多め & 難しめ）

対象：計算機科学コース・ソフトウェア科学コース・数理科学コース

担当教員：大木 健太郎

2013 年 7 月 18 日 (木曜日 5 時限)

1. 区間 $[-1,1]$ において，次の関数を最大にする $x = x^* \in [-1,1]$ を求めたい．

$$S(x) = -x^4 + 2x^3 - x^2 - x + 1$$

   (a) $p(x) = \frac{dS(x)}{dx}$ が $[-1,1]$ の間に根を 1 つ持つことを，Sturm の定理を用いて示せ．

   (b) $p(x) = 0$ の解を Newton 法で求めたい．10 進 2 桁の浮動小数点数表示の計算機を用いることを想定し，初期値を $x^{(0)} := -1$ として，2 回反復した値 $x^{(2)}$ を求めよ．

2. Hermite 行列 $A \in \mathbb{C}^{n\times n}$ の固有値問題を考える．

   (a) $n = 2$ のとき，Hermite 行列 A を対角化するユニタリ行列を求めたい．ユニタリ行列

$$P = \begin{bmatrix} 1 & 0 \\ 0 & e^{\mathrm{i}\phi} \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \theta, \phi \in [0, 2\pi) \tag{1}$$

   を用いて，P^*AP を対角行列にできることを示せ．ただし，P^* は行列 P の Hermite 共役を表す．

   (b) 式 (1) で定義されるユニタリ行列 P_k を用いて Hermite 行列を対角化することを考える．ただし，$p, q = 1, 2, \cdots n$ とし，pq 成分を消去するように定義されたユニタリ行列

$$P_k = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos(\theta_k) & \cdots & \sin(\theta_k) & & \\ & & \vdots & \ddots & \vdots & & \\ & & -e^{\mathrm{i}\phi_k}\sin(\theta_k) & \cdots & e^{\mathrm{i}\phi_k}\cos(\theta_k) & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, \quad \theta_k, \phi_k \in [0, 2\pi) \tag{2}$$

   を用いるものとする．ϕ_k, θ_k を適切に選び，非対角要素を減らすことを考える．$A^{(0)} := A$ とし，$A^{(k)} := P_k^* A^{(k-1)} P_k$ とする．このとき，非ゼロの非対角要素を消去し続けることで，$F_k := \sum_{i\neq j} |A_{ij}^{(k)}|^2 \to 0$，$k \to \infty$ とできることを示せ．

(c) (b) の方法で次の行列を対角化したい．

$$A = \begin{bmatrix} 1 & 1+3\mathrm{i} & -2\mathrm{i} \\ 1-3\mathrm{i} & 1 & 2 \\ 2\mathrm{i} & 2 & 1 \end{bmatrix}, \quad \mathrm{i} := \sqrt{-1}$$

$|F_{k+1} - F_k|$ を最も大きくする方法で，$k = 1$ まで計算せよ．

3. 数値計算問題の解きにくさについて，次の問に答えよ．

(a) 次の連立一次方程式は，数値計算的に解きにくい．その理由を述べよ．

$$\begin{bmatrix} 2 & -1 \\ -2000 & 999.9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(b) 関数の集合 X には内積 $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ が定義されているものし，$\{\phi_j\}_j \subseteq X$ は基底であるとする．関数 $f \in X$ を，$n+1$ 個の基底の線形和

$$f_n = \sum_{j=0}^{n} c_j \phi_j$$

で，最小二乗近似する問題を考える．

$$\min_{c_j,\ j=0,\cdots,n} \langle f - f_n, f - f_n \rangle$$

$\{\phi_j\}_j$ が直交基底である場合と非直交基底を用いた場合に，どちらが解きにくいか，理由を付けて答えよ．

4. 次の行列の最大特異値を求めよ．数値計算的に求める場合は，アルゴリズムも明記すること．

$$X = \begin{bmatrix} 1 & 2 & 3 & 1 \\ -1 & 1 & 0 & -1 \end{bmatrix} \tag{3}$$

解答例

1. (a) $p(x)$ は

$$p(x) = -4x^3 + 6x^2 - 2x - 1 \tag{4}$$

である．Sturm 列は，$p_0(x)$ として

$$p_1(x) = \frac{d}{dx}p(x) = -12x^2 + 12x - 2, \tag{5}$$
$$p_2(x) = g_1(x)p_1(x) - p_0(x),\ g_1(x) = a_1x + b_1 \tag{6}$$
$$p_3(x) = g_2(x)p_2(x) - p_1(x),\ g_2(x) = a_2x + b_2 \tag{7}$$

である．$g_1(x)$，$g_2(x)$ の係数をそれぞれ求めると，

$$a_1 = \frac{1}{3},\ b_1 = -\frac{1}{6},\ a_2 = 18,\ b_2 = 18$$

より，

$$p_2(x) = -\frac{2}{3}(x-2), \tag{8}$$
$$p_3(x) = 26 \tag{9}$$

を得る．考えている閉区間は $[-1, 1]$ なので，

$$\begin{pmatrix} p(x) & p_1(x) & p_2(x) & p_3(x) \end{pmatrix}$$

にそれぞれ端の値を代入して符号変化を調べればよい．

$$\begin{pmatrix} p(-1) & p_1(-1) & p_2(-1) & p_3(-1) \end{pmatrix} = \begin{pmatrix} 11 & -26 & 2 & 26 \end{pmatrix}, \tag{10}$$
$$\begin{pmatrix} p(1) & p_1(1) & p_2(1) & p_3(1) \end{pmatrix} = \begin{pmatrix} -1 & -2 & \frac{2}{3} & 26 \end{pmatrix} \tag{11}$$

より，符号変化の回数 $N(x)$ の差は

$$n = N(-1) - N(1) = 2 - 1 - 1 \tag{12}$$

となる．したがって，$[-1, 1]$ の間にある $p(x) = 0$ の解は唯一つ存在する．

(b) Newton 法より，

$$x^{(k+1)} = x^{(k)} - \frac{p(x^{(k)})}{p'(x^{(k)})} = x^{(k)} - \frac{p(x^{(k)})}{p_1(x^{(k)})} \tag{13}$$

を解けばよい．$x^{(0)} = -1$ より，

$$\begin{aligned} x^{(1)} &= x^{(0)} - \frac{p(-1)}{p_1(-1)} = -1 + \frac{11}{26} \\ &\simeq -1 + 0.42 = -0.58 \end{aligned} \tag{14}$$

を得る．$0.58^2 = 0.3364 \simeq 0.34$，$0.58^3 = 0.195112 \simeq 0.20$ であることを用いると，

$$p(-0.58) = -4 \times (-0.58)^3 + 6 \times (-0.58)^2 - 2 \times (-0.58) - 1 \simeq 3.0$$
$$p_1(-0.58) = -12 \times (-0.58)^2 + 12 \times (-0.58) - 2 \simeq -4.1 - 7.0 - 2.0 \simeq -13$$

である．これを用いて同様に計算すると，

$$x^{(2)} = x^{(1)} - \frac{p(-0.58)}{p_1(-0.58)} \simeq -0.58 + \frac{3}{13} \simeq -0.35 \tag{15}$$

を得る．

2. (a)

$$A = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix}, \quad a, c \in \mathbb{R},\ b \in \mathbb{C} \tag{16}$$

とする．このとき，$b = |b|e^{\mathrm{i}\theta_b}$ とおき，$\phi = -\theta_b$ と置くと，

$$P^* A P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a & |b| \\ |b| & c \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{17}$$

となり，実対称行列の対角化問題となる．θ を

$$\theta = \frac{1}{2}\arctan\left(\frac{-2|b|}{a-c}\right) \tag{18}$$

とすれば，式 (1) で定義されるユニタリ行列で対角化可能である．

(b) 簡単のために，最初のステップのみを示す．$B = (B_{ij}) = A^{(1)}$ とし，P は A の (p, q) 成分を消去するようにパラメータが選ばれているものとする；$q > p$ とすると，

$$P_{ij} = \begin{cases} 1, & i = j,\ i \neq p, q \\ \cos(\theta), & i = j = p \\ \sin(\theta), & i = p,\ j = q \\ -e^{\mathrm{i}\phi}\sin(\theta), & i = q,\ j = p \\ e^{\mathrm{i}\phi}\cos(\theta), & i = j = q \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

である．このとき，$j \neq p, q$ とすると

$$\begin{aligned} B_{pj} &= \sum_{k,l=1}^{n} P^*_{pk} A_{kl} P_{lj} = \sum_{k,l=1}^{n} \bar{P}_{kp} A_{kl} P_{lj} \\ &= \bar{P}_{pp} A_{pj} + \bar{P}_{qp} A_{qj} = \cos(\theta) A_{pj} - e^{-\mathrm{i}\phi}\sin(\theta) A_{qj}, \end{aligned} \tag{20}$$

$$\begin{aligned} B_{qj} &= \sum_{k,l=1}^{n} P^*_{qk} A_{kl} P_{lj} = \sum_{k,l=1}^{n} \bar{P}_{kq} A_{kl} P_{lj} \\ &= \bar{P}_{qp} A_{pj} + \bar{P}_{qq} A_{qj} = \sin(\theta) A_{pj} + e^{-\mathrm{i}\phi}\cos(\theta) A_{qj} \end{aligned} \tag{21}$$

となる．他の要素は，次のようになる．

$$
\begin{aligned}
B_{ij} =& A_{ij}, \quad i, j \neq p, q && (22)\\
B_{pp} =& \bar{P}_{pp}A_{pp}P_{pp} + \bar{P}_{pp}A_{pq}P_{qp} + \bar{P}_{qp}A_{qp}P_{pp} + \bar{P}_{qp}A_{qq}P_{qp}\\
=& \cos^2(\theta)A_{pp} + \sin^2(\theta)A_{qq} - e^{-\mathrm{i}\phi}\sin(\theta)\cos(\theta)A_{qp} - e^{+\mathrm{i}\phi}\sin(\theta)\cos(\theta)A_{pq}\\
=& \frac{A_{pp}+A_{qq}}{2} + \frac{A_{pp}-A_{qq}}{2}\cos(2\theta) - \mathrm{Re}(e^{-\mathrm{i}\phi}A_{qp})\sin(2\theta) && (23)\\
B_{qq} =& \bar{P}_{pq}A_{pp}P_{pq} + \bar{P}_{pq}A_{pq}P_{qq} + \bar{P}_{qq}A_{qp}P_{pq} + \bar{P}_{qq}A_{qq}P_{qq}\\
=& \sin^2(\theta)A_{pp} + \cos^2(\theta)A_{qq} + e^{-\mathrm{i}\phi}\sin(\theta)\cos(\theta)A_{qp} + e^{+\mathrm{i}\phi}\sin(\theta)\cos(\theta)A_{pq}\\
=& \frac{A_{pp}+A_{qq}}{2} - \frac{A_{pp}-A_{qq}}{2}\cos(2\theta) + \mathrm{Re}(e^{-\mathrm{i}\phi}A_{qp})\sin(2\theta) && (24)\\
B_{pq} =& \bar{P}_{pp}A_{pp}P_{pq} + \bar{P}_{pp}A_{pq}P_{qq} + \bar{P}_{qp}A_{qp}P_{pq} + \bar{P}_{qp}A_{qq}P_{qq} = 0\\
=& \cos(\theta)\sin(\theta)(A_{pp}-A_{qq}) + e^{\mathrm{i}\phi}\cos^2(\theta)A_{pq} - e^{-\mathrm{i}\phi}\sin^2(\theta)A_{qp}\\
=& \frac{A_{pp}-A_{qq}}{2}\sin(2\phi) - \mathrm{i}\mathrm{Im}(e^{\mathrm{i}\phi}A_{pq}) + \mathrm{Re}(e^{\mathrm{i}\phi}A_{pq})\cos(2\theta) && (25)
\end{aligned}
$$

ただし，複素数 α に対して，

$$\mathrm{Re}(\alpha) := \frac{\alpha+\bar{\alpha}}{2}, \quad \mathrm{Im}(\alpha) := \frac{\alpha-\bar{\alpha}}{2\mathrm{i}}$$

を表す．ϕ は $\mathrm{Im}(e^{\mathrm{i}\phi}A_{pq}) = 0$ となるように選んでいるので，教科書の p. 93 の議論の通り，

$$F_1 = F_0 - 2|A_{pq}|^2 \tag{26}$$

となる．同様の議論を繰り返して，任意の k に対して

$$F_{k+1} - F_k = -2|A_{qp}^{(k)}|^2 \leq 0 \tag{27}$$

を得るので，各ステップ k で非ゼロの非対角成分を選択することで，漸近的に非対角成分をゼロにすることができる．その結果，対角成分には固有値が残る．

(c) 最初のステップにおいて，非対角要素で最も大きな成分は，$|A_{2,1}| = \sqrt{10}$ である．$A_{2,1}$ の実部は 1，虚部は -3 であるので，$\phi_1 = -\arctan(-3)$ とおくと，

$$e^{-\mathrm{i}\phi_1} = \underbrace{\cos(\phi_1)}_{=\frac{1}{\sqrt{10}}} + \mathrm{i}\underbrace{\sin(\phi_1)}_{=\frac{3}{\sqrt{10}}}$$

であるので，

$$\cos(2\theta_1) = \frac{|A_{pp}-A_{qq}|}{\sqrt{4|A_{pq}|^2 + (A_{pp}-A_{qq})^2}}, \ \cos(\theta_1) = \sqrt{\frac{1}{2}(1+\cos(2\theta_1))},$$

$$\sin(\theta_1) = \mathrm{sgn}(-A_{pq}(A_{pp}-A_{qq}))\sqrt{\frac{1}{2}(1-\cos(2\theta_1))}$$

を用いると，

$$
\begin{aligned}
A^{(1)} &= P_1^* A P_1 \\
&= \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \sqrt{10} & -2\mathrm{i} \\ \sqrt{10} & 1 & 2e^{-\mathrm{i}\phi_1} \\ 2\mathrm{i} & 2e^{\mathrm{i}\phi_1} & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) & 0 \\ -\sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1-\sqrt{10}\sin(2\theta) & 0 & -2\mathrm{i}\cos(\theta_1) - e^{-\mathrm{i}\phi_1}2\sin(\theta_1) \\ 0 & 1+\sqrt{10}\sin(2\theta_1) & -2\mathrm{i}\sin(\theta_1) + e^{-\mathrm{i}\phi_1}2\cos(\theta_1) \\ 2\mathrm{i}\cos(\theta_1) - e^{\mathrm{i}\phi_1}2\sin(\theta_1) & 2\mathrm{i}\sin(\theta_1) + e^{\mathrm{i}\phi_1}2\cos(\theta_1) & 1 \end{bmatrix}
\end{aligned}
$$

解答はここまで書いてあればよく，各要素を具体的に求める必要はない．

3. (a) 行列 A は，1 行目と 2 行目のベクトルは，独立であるものの従属性が大きい．そのため，行列 A の条件数が悪くなり，浮動小数点数表示によって誤差が大きくなりやすい．

   (b) 基底 $\{\phi_j\}_j$ を用いたとき，

$$\langle f - f_n, f - f_n \rangle = (Ac - b)^\top A^{-1}(Ac - b) + \langle f, f \rangle - b^\top A^{-1} b \tag{28}$$

   となる．ここで

$$A = (A_{ij}) = (\langle \phi_i, \phi_j \rangle),\ c = (c_j),\ b = (b_j) = (\langle \phi_j, f \rangle) \tag{29}$$

   とした．A が逆行列をもち，かつ正定値対称行列であることに注意すると，

$$Ac - b = 0 \tag{30}$$

   を満たす $c \in \mathbb{R}^{n+1}$ が最適解である．
   ここで，$\{\phi_j\}_j$ が直交基底であるならば，A は対角行列となり，$c = A^{-1}b$ が少ない計算で求められる．一般の場合は計算量が $O(n^3)$ 程度になってしまい，数値誤差が大きくなる可能性がある．したがって，直交基底を用いたほうが解きやすい．

4. 行列 X の最大特異値は，$X^\top X$ の非ゼロの最大固有値の平方根，または $XX^\top$ の非ゼロの最大固有値の平方根である．数値計算的には小さな次元で計算することが得なので，$XX^\top$ を求めると，

$$XX^\top = \begin{bmatrix} 15 & 0 \\ 0 & 3 \end{bmatrix} \tag{31}$$

を得る．この行列はすでに対角行列であるので，対角要素の平方根が特異値である．よって，X の最大特異値は，$\sqrt{15}$.