Linear Algebra II

07/04/2014, Monday, 9:00-12:00

1 $(10 + 5 = 15 \text{ pts})$ **Gram-Schmidt process**

Consider the vector space $\mathbb{R}^4$ with the inner product

$$\langle x, y\rangle = x^T y.$$

Let $S \subset \mathbb{R}^4$ be the subspace given by

$$S = \operatorname{span}\left\{ \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\-1\end{bmatrix} \right\}.$$

(a) Apply the Gram-Schmidt process to obtain an orthonormal basis for S.

(b) Find the closest element in the subspace S to the vector

$$\begin{bmatrix}a\\b\\b\\a\end{bmatrix}$$

where a and b are real numbers.

Required Knowledge: **inner product, Gram-Schmidt process, least squares**

Solution:

(1a): Let

$$x_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad x_2 = \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \quad \text{and} \quad x_3 = \begin{bmatrix}1\\0\\0\\-1\end{bmatrix}.$$

By applying the Gram-Schmidt process, we obtain:

$$u_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$u_2 = \frac{x_2 - p_1}{\|x_2 - p_1\|} \qquad p_1 = \langle x_2, u_1\rangle \cdot u_1$$

$$= \frac{1}{4}\cdot 2\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$x_2 - p_1 = \begin{bmatrix}1\\1\\0\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$$

$$\|x_2 - p_1\|^2 = \frac{1}{4}\cdot 4 = 1$$

$$u_2 = \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$$

$$u_3 = \frac{x_3 - p_2}{\|x_3 - p_2\|} \qquad p_2 = \langle x_3, u_1\rangle \cdot u_1 + \langle x_3, u_2\rangle \cdot u_2$$

$$= \frac{1}{4}\cdot 2\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$$

$$x_3 - p_2 = \begin{bmatrix}1\\0\\0\\-1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}$$

$$\|x_3 - p_2\|^2 = \frac{1}{4}\cdot 4 = 1$$

$$u_3 = \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}.$$

(1b): The closest element in S to x can be found by projection:

$$p = \langle x, u_1\rangle \cdot u_1 + \langle x, u_2\rangle \cdot u_2 + \langle x, u_3\rangle \cdot u_3.$$

Thus, we have

$$p = \frac{1}{4}(a+b+b+a)\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{1}{4}(a+b-b-a)\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} + \frac{1}{4}(a-b+b-a)\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix} = \frac{1}{2}(a+b)\begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

Consider the matrix

$$M = \begin{bmatrix} 9 & -6 \\ 5 & -3 \end{bmatrix}.$$

By using the Cayley-Hamilton theorem, find a and b such that

$$M^5 = aM + bI.$$

REQUIRED KNOWLEDGE: **Cayley-Hamilton theorem**

SOLUTION:

The characteristic polynomial of M is given by

$$\det(M - \lambda I) = \det\left(\begin{bmatrix} 9-\lambda & -6 \\ 5 & -3-\lambda \end{bmatrix}\right) = (9-\lambda)(-3-\lambda) + 30 = \lambda^2 - 6\lambda + 3.$$

It follows from Cayley-Hamilton theorem that

$$M^2 - 6M + 3I = 0.$$

Then, we have

$$M^2 = 6M - 3I = 3(2M - I).$$

This results in

$$M^3 = MM^2 = 3M(2M - I) = 3(2M^2 - M) = 3\big(2(6M - 3I) - M\big) = 3(11M - 6I).$$

Therefore, we obtain

$$\begin{aligned} M^5 &= M^2M^3 = 9(2M - I)(11M - 6I) \\ &= 9(22M^2 - 23M + 6I) \\ &= 9\big(66(2M - I) - 23M + 6I\big) \\ &= 9(132M - 66I - 23M + 6I) = 9(109M - 60I) = 981M - 540I. \end{aligned}$$

As such, $a = 981$ and $b = -540$.

Consider the matrix

$$M = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 3 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{bmatrix}.$$

(a) Find the singular values of M.

(b) Find a singular value decomposition for M.

(c) Find the best rank 2 approximation of M.

REQUIRED KNOWLEDGE: **singular value decomposition, lower rank approximations**

SOLUTION:

(3a): Note that

$$M^T M = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 36 \end{bmatrix}.$$

Then, the eigenvalues of $M^T M$ are given by

$$\lambda_1 = 36, \quad \lambda_2 = 16, \quad \text{and} \quad \lambda_3 = 4$$

and hence the singular values by

$$\sigma_1 = 6, \quad \sigma_2 = 4, \quad \text{and} \quad \sigma_3 = 2.$$

(3b): Three eigenvectors of $M^T M$ corresponding to the eigenvalues λ_1, λ_2, and λ_3 can be given by

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

As such, we have

$$V = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Note that the rank of M is equal to the number of nonzero singular values. Thus, $r = \operatorname{rank}(M) = 3$. By using the formula

$$u_i = \frac{1}{\sigma_i} M v_i$$

for $i = 1, 2, 3$, we obtain

$$u_1 = \frac{1}{6} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 3 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

$$u_2 = \frac{1}{4} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 3 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

$$u_3 = \frac{1}{2} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 3 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}.$$

The last column vector of the matrix U can be found by looking at the null space of M^T:

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 2 & -2 & 2 \\ 3 & 3 & 3 & -3 \end{bmatrix} y = 0.$$

By row operations, we obtain

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} y = 0.$$

This yields, for instance,

$$y = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Thus, we get

$$u_3 = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Finally, the SVD can be given by:

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 3 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

(3c): The best rank 2 approximation can be obtained as follows:

$$\begin{aligned} \tilde{M} &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 & 0 \\ 3 & 2 & 0 \\ 3 & -2 & 0 \\ -3 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & -2 & 3 \\ 0 & 2 & -3 \end{bmatrix}. \end{aligned}$$

(a) Let A be a square matrix. Show that

$$\det(e^A) = e^{\mathrm{tr}(A)}.$$

(b) Let B be an orthogonal matrix. Find out the singular values of B.

(c) Let C and D be $n \times n$ matrices. Suppose that C is orthogonal. Find out the relationship between the singular values of D and those of CD.

REQUIRED KNOWLEDGE: **eigenvalues, orthogonal matrices and singular values**

SOLUTION:

(4a): Let (λ, x) be an eigenpair of A, that is

$$Ax = \lambda x.$$

Note that

$$A^k x = \lambda^k x.$$

Thus, we have

$$e^A x = (I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots)x = (1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots)x = e^\lambda x.$$

In other words, if λ is an eigenvalue of A then e^λ is an eigenvalue of e^A. Since the determinant of a matrix equals to the product of eigenvalues, we have

$$\det(e^A) = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n}$$

where λ_i for $i = 1, 2, \ldots, n$ are the eigenvalues of A. Hence, we have

$$\det(e^A) = e^{\lambda_1 + \lambda_2 + \cdots + \lambda_n}.$$

Since the sum of the eigenvalues of a matrix equals to its trace, we get

$$\det(e^A) = e^{\lambda_1 + \lambda_2 + \cdots + \lambda_n} = e^{\mathrm{tr}(A)}.$$

(4b): Approach 1: The singular values of B are the square roots of the eigenvalues of $B^T B$. Since B is orthogonal, $B^T B = I$. Hence, all singular values of B are equal to 1.

Approach 2: Since B is an orthogonal matrix, we have the SVD

$$B = U \Sigma V^T$$

where $U = B$, and $\Sigma = V = I$. As such, singular values of B are all 1.

(4c): Approach 1: The singular values of CD are the square roots of the eigenvalues of $(CD)^T CD$. Note that

$$(CD)^T CD = D^T C^T CD = D^T D$$

where the last equality follows from the fact that C is orthogonal. As such, the singular values of CD and D are the same.

Approach 2: Let

$$D = U \Sigma V^T$$

be an SVD of D. Then, we have

$$CD = CU \Sigma V^T. \qquad (\star)$$

Since both C and U are orthogonal, so is their product CD. Thus, $(\star)$ is an SVD for CD. Consequently, D and CD have the same singular values.

(a) Consider the function

$$f(x, y) = 6xy^2 - 2x^3 - 3y^4.$$

Find the stationary points of f and determine whether its stationary points are local minimum/maximum or saddle points.

(b) Let

$$M = \begin{bmatrix} 2 & 1 & a \\ 1 & 2 & 1 \\ a & 1 & 2 \end{bmatrix}$$

where a is a real number. Determine all values of a for which M is

(i) positive definite.

(ii) negative definite.

REQUIRED KNOWLEDGE: **stationary points, positive definiteness**

SOLUTION:

(5a): In order to find the stationary points, we need the partial derivatives:

$$f_x = 6y^2 - 6x^2 \quad \text{and} \quad f_y = 12xy - 12y^3.$$

Then, $(\bar{x}, \bar{y})$ is a stationary point if and only if

$$\bar{y}^2 - \bar{x}^2 = 0$$
$$\bar{x}\bar{y} - \bar{y}^3 = 0.$$

The second yields $\bar{y} = 0$ or $\bar{x} = \bar{y}^2$. If $\bar{y} = 0$, then we get from the first $\bar{x} = 0$. Hence,

$$(\bar{x}_1, \bar{y}_1) = (0, 0)$$

is a stationary point. If $\bar{x} = \bar{y}^2$, then we get $\bar{y}^4 = \bar{y}^2$ from the first. This holds if and only if $\bar{y} \in \{-1, 0, 1\}$. Thus, we obtain two more stationary points

$$(\bar{x}_2, \bar{y}_2) = (1, -1) \quad \text{and} \quad (\bar{x}_3, \bar{y}_3) = (1, 1).$$

To determine the character of these points, we need the second order partial derivatives:

$$f_{xx} = -12x, \quad f_{xy} = 12y, \quad \text{and} \quad f_{yy} = 12x - 36y^2.$$

For the stationary point $(\bar{x}_1, \bar{y}_1) = (0, 0)$, we have

$$H_{(0,0)} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}_{(0,0)} = \begin{bmatrix} 12x & 12y \\ 12y & 12x - 36y^2 \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since this matrix has only zero eigenvalues, we cannot determine the nature of the corresponding stationary point.

For the stationary point $(\bar{x}_2, \bar{y}_2) = (1, -1)$, we have

$$H_{(1,-1)} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}_{(1,-1)} = \begin{bmatrix} -12x & 12y \\ 12y & 12x - 36y^2 \end{bmatrix}_{(1,-1)} = \begin{bmatrix} -12 & -12 \\ -12 & -24 \end{bmatrix} = -12 \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Note that the characteristic equation for the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ is given by

$$(\lambda - 1)(\lambda - 2) - 1 = \lambda^2 - 3\lambda 1 = 0.$$

Thus, we find the roots as

$$\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}.$$

Note that both these numbers are positive. Since $-12\lambda_{1,2}$ are the eigenvalues of the Hessian, it is negative definite. This means the corresponding stationary point is a local maximum.

For the stationary point $(\bar{x}_3, \bar{y}_3) = (1, 1)$, we have

$$H_{(1,1)} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}_{(1,1)} = \begin{bmatrix} -12x & 12y \\ 12y & 12x - 36y^2 \end{bmatrix}_{(1,1)} = \begin{bmatrix} -12 & 12 \\ 12 & -24 \end{bmatrix} = -12 \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Note that the characteristic equation for the matrix $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ is exactly the same as for the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. As such, $H_{(1,1)}$ is negative definite. Therefore, the corresponding stationary point is a local maximum.

(5b)(i): A symmetric matrix is positive definite if and only if all its leading principal minors are positive. Note that the leading principal minors of M are given by:

$$\det(2) = 2, \quad \det\left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right) = 4-1 = 3, \quad \text{and} \quad \det\left(\begin{bmatrix} 2 & 1 & a \\ 1 & 2 & 1 \\ a & 1 & 2 \end{bmatrix}\right) = 8+a+a-2a^2-2-2 = 4+2a-2a^2.$$

Then, the matrix M is positive definite if and only if

$$a^2 - a - 2 < 0.$$

Since $a^2-a-2 = (a+1)(a-2)$, we can conclude that M is positive definite if and only if $-1 < a < 2$.

(5b)(ii): The matrix M is negative definite if and only if $-M$ is positive definite. Since the first leading principal minor of $-M$ is -2, there are no values of a and b rendering M negative definite.

Consider the matrix

$$M = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

(a) Find the eigenvalues of M.

(b) Is M diagonalizable? Why?

(c) Put M into the Jordan canonical form.

REQUIRED KNOWLEDGE: **eigenvalues/vectors, diagonalization, Jordan canonical form**

SOLUTION:

(6a): Charateristic polynomial of M can be found as

$$\begin{aligned} \det(M-\lambda) &= \det\left(\begin{bmatrix} -1-\lambda & 1 & 0 \\ -1 & -\lambda & 1 \\ -1 & 0 & 1-\lambda \end{bmatrix} \right) \\ &= \lambda(1+\lambda)(1-\lambda) - 1 + 1 - \lambda \\ &= \lambda(1-\lambda^2) - \lambda \\ &= \lambda(1-\lambda^2-1) = -\lambda^3. \end{aligned}$$

Therefore, M has only zero eigenvalues.

(6b): The matrix M is diagonalizable if and only if it has 3 linearly independent eigenvectors. To find the eigenvectors, we need to solve the equation $Mx = 0$ since eigenvalues are all zero. Note that the system of equations

$$\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} x = 0$$

is equivalent to that of

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} x = 0.$$

Therefore, the general solution is of the form

$$x = \begin{bmatrix} a \\ a \\ a \end{bmatrix}$$

where a is a scalar. This means that we can find only one linearly dependent eigenvector for the zero eigenvalue. Consequently, M is not diagonalizable.

(6c): Since there is only one linearly independent eigenvector, Jordan canonical form consists of one block. Note that

$$M^2 = \begin{bmatrix} 0 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{and} \quad M^3 = 0.$$

Next, we solve

$$M^2 v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

One possible solution is

$$v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Note that

$$Mv = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Let

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

and note that

$$\underbrace{\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}}_{T} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{J}.$$