تمرین شماره سوم درس یادگیری ماشین

پریا پاسه‌ورز

شماره دانشجویی: 810101393

سوال سوم)

الف) رابطه تجزیه یک ماتریس مربعی را به کمک روش eigen decomposition بیان کنید.

برای هر ماتریس مربعی مثل A، رابطه eigen decomposition به شرح زیر است:

$$A = Q\Lambda Q^{-1}$$

در اینجا، Q ماتریسی است که از کنار هم قرار دادن ستونی مقادیر بردارهای ویژه ماتریس A به دست آمده است.

ماتریس Λ نیز یک ماتریس قطری است که مقادیر روی قطرش مقادیر ویژه ماتریس A هستند.

دقت کنید این روش فقط زمانی جواب می‌دهد که بردارهای ویژه از هم مستقل باشند، در غیر این صورت باید از روش Jordan Decomposition استفاده کنیم.

مثال:

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$$

$$det\left(\begin{bmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{bmatrix}\right) = 0$$

$$(4-\lambda)(3-\lambda) - 2 \cdot 1 = 0$$

$$\lambda^2 - 7\lambda + 10 = 0$$

$$(\lambda - 5)(\lambda - 2) = 0$$

$$\lambda_1 = 5, \quad \lambda_2 = 2$$

$$(A - 5I)v = 0 \;\Rightarrow\; \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} v = 0 \;\Rightarrow\; v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$(A - 2I)v = 0 \;\Rightarrow\; \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} v = 0 \;\Rightarrow\; v = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \quad \Rightarrow \quad Q^{-1} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$$

ب) یکی از روش‌های تصحیح شده نیوتون را که از تجزیه مقادیر ویژه ماتریس هسین برای تقریب این ماتریس استفاده می‌کند بیان کنید.

رابطه اساسی الگوریتم نیوتون را در نظر بگیرید:

$$x_{k+1} = x_k - H_k^{-1}\nabla f(x_k)$$

با استفاده از eigen decomposition، ماتریس هسین را تجزیه می‌کنیم:

$$H_k = Q\Lambda Q^{-1}$$

برای اطمینان از positive definite بودن ماتریس هسین، مقدار ویژه‌های منفی یا خیلی کوچک موجود در ماتریس Λ را با یک مقدار مثبت کوچک جایگزین می‌کنیم.

ب) تابع زیر را در نظر بگیرید:

$$\boldsymbol{f(x_1, x_2) = 4x_1^3 + 3x_1x_2 + 5x_2^2 + 2x_1^2x_2}$$

با نقطه اولیه (1, 0)، یک پله از الگوریتم بالا را با طول پله 1 بردارید.

اول گرادیان تابع f را محاسبه می‌کنیم:

$$\frac{\partial f}{\partial x_1} = 12x_1^2 + 3x_2 + 4x_1x_2$$

$$\frac{\partial f}{\partial x_2} = 3x_1 + 10x_2 + 2x_1^2$$

$$\nabla f(x_1, x_2) = [\,12x_1^2 + 3x_2 + 4x_1x_2\,, 3x_1 + 10x_2 + 2x_1^2]$$

$$\nabla f(1,0) = [\,12\,, 5]$$

سپس ماتریس هسین را محاسبه می‌کنیم:

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{\partial}{\partial x_1}(12x_1^2 + 3x_2 + 4x_1x_2) = 24x_1 + 4x_2$$

$$\frac{\partial^2 f}{\partial x_2^2} = \frac{\partial}{\partial x_2}(3x_1 + 10x_2 + 2x_1^2) = 10$$

$$\frac{\partial^2 f}{\partial x_1\,\partial x_2} = \frac{\partial}{\partial x_2}(12x_1^2 + 3x_2 + 4x_1x_2) = 3 + 4x_1$$

$$\frac{\partial^2 f}{\partial x_2\,\partial x_1} = 3 + 4x_1$$

$$H(x_1, x_2) = \begin{bmatrix} 24x_1 + 4x_2 & 3 + 4x_1 \\ 3 + 4x_1 & 10 \end{bmatrix}$$

$$H(1,0) = \begin{bmatrix} 24 & 7 \\ 7 & 10 \end{bmatrix}$$

حال مقادیر ویژه این ماتریس را پیدا می‌کنیم:

$$\det(H - \lambda I) = 0$$

$$det\left(\begin{bmatrix} 24-\lambda & 7 \\ 7 & 10-\lambda \end{bmatrix}\right) = 0$$

$$(24-\lambda)(10-\lambda) - 7 \times 7 = 0$$

$$\lambda^2 - 34\lambda + 191 = 0$$

$$\lambda = \frac{34 \pm \sqrt{1156 - 764}}{2}$$

$$\lambda = \frac{36 \pm 18.44}{2}$$

$$\lambda_1 = 26.89, \quad \lambda_2 = 7.10$$

سپس بردارهای ویژه آن را:

$$(A - 26.89I)v = 0 \;\Rightarrow\; \begin{bmatrix} 24-26.89 & 7 \\ 7 & 10-26.89 \end{bmatrix} v = 0 \;\Rightarrow\; v = \begin{bmatrix} 2.41 \\ 1 \end{bmatrix}$$

$$(A - 7.10I)v = 0 \;\Rightarrow\; \begin{bmatrix} 24-7.10 & 7 \\ 7 & 10-7.10 \end{bmatrix} v = 0 \;\Rightarrow\; v = \begin{bmatrix} -0.41 \\ 1 \end{bmatrix}$$

$$Q = \begin{bmatrix} 2.4 & -0.41 \\ 1 & 1 \end{bmatrix} \quad\Rightarrow\quad Q^{-1} = \frac{1}{2.81}\begin{bmatrix} 1 & 0.41 \\ -1 & 2.4 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 26.89 & 0 \\ 0 & 7.10 \end{bmatrix}$$

$$H = \begin{bmatrix} 2.4 & -0.41 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 26.89 & 0 \\ 0 & 7.10 \end{bmatrix}\frac{1}{2.81}\begin{bmatrix} 1 & 0.41 \\ -1 & 2.4 \end{bmatrix} = \begin{bmatrix} 21.93 & 6.92 \\ 7.04 & 9.98 \end{bmatrix}$$

$$x_{k+1} = x_k - H_k^{-1}\nabla f(x_k)$$

$$x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 21.93 & 6.92 \\ 7.04 & 9.98 \end{bmatrix}\right)^{-1}\begin{bmatrix} 12 \\ 5 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1.14 \end{bmatrix}$$