Stats 230, Homework 2

Due date: **February 4**

1. (Lange Exercise 7.6) Find by hand the Cholesky decomposition of the matrix

$$A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

2. (Lange Exercise 7.8) Suppose the matrix $A = \{a_{ij}\}$ is banded in the sense that $a_{ij} = 0$ when $|i - j| > d$. Prove that the Cholesky decomposition $B = \{b_{ij}\}$ also satisfies the band condition $b_{ij} = 0$ when $|i - j| > d$.

   Small Bonus: what can we tell about sparse matrices in general?

3. (Lange Exercise 7.11) If $X = QR$ is the QR decomposition of X, where X has linearly independent columns, then show that the projection matrix

$$X(X^T X)^{-1} X^T = QQ^T.$$

   In addition, show that $|\det(X)| = |\det(R)|$ when X is square and in general that $\det(X^T X) = [\det(R)]^2$.

4. (Lange Exercise 8.4, modified) Show that the reflection matrix

$$A = \begin{pmatrix} cos(\theta) & sin(\theta) \\ sin(\theta) & -cos(\theta) \end{pmatrix}$$

   is orthogonal and find its eigenvalues and eigenvectors.

5. (Lange Exercise 8.5, modified) Suppose λ is an eigenvalue of the orthogonal matrix O with corresponding eigenvector $\mathbf{v}$. Show that if $\mathbf{v}$ has real valued entries, then $\lambda = \pm 1$.

   Hint: use properties of the norms of orthogonal matrices.

6. (Lange Exercise 9.3, modified) It can be shown that the matrix norm induced by the Euclidean (L_2) norm of matrix A is equal to the largest singular value of this matrix. Now, let A be an invertible $m \times m$ matrix with singular values $\sigma_1, \cdots, \sigma_m$. Recall that the L_2 condition number is defined as $\text{cond}_2(A) = ||A||_2 ||A^{-1}||_2$. Prove that

$$\text{cond}_2(A) = \frac{\max_i \sigma_i}{\min_i \sigma_i}.$$

7. Simulation of multivariate normal random vectors.

(a) Write an R function that takes as input an n dimensional numeric vector $\boldsymbol{\mu}$ and a $n \times n$ positive definite matrix $\boldsymbol{\Sigma}$ and returns N realizations from the multivariate normal distribution $\mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, using Cholesky decomposition.

(b) Document this function and add it to your package

(c) Create a test case with $n = 4$ and $N = 100$ and use sample mean and sample covariance matrices to (somewhat informally) validate your function.

8. Download the file `homework2_regression.csv`. The file contains simulated data with a response vector and 5 covariates, including a dummy one for the intercept. In the tasks below, pay attention to the order of operations, so that your computations are performed efficiently.

   (a) Obtain OLS estimates of the regression coefficients using a QR decomposition (implement in a function, document, and add to the package)

   (b) Obtain OLS estimates of the regression coefficients using the SVD decomposition (implement in a function, document, and add to the package)

   (c) Benchmark computational efficiency of both implementations and comment on your results

Hint: you may find it helpful to look at R snippets when working on the last two problems: `https://r-snippets.readthedocs.io/en/latest/la/index.html`

1. (Lange Exercise 7.6) Find by hand the Cholesky decomposition of the matrix

$$A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

$$\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} = LL^T$$

$$= \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} \\ 0 & l_{22} \end{bmatrix}$$

$$\Rightarrow a_{11} = l_{11}^2 \quad \Rightarrow \quad l_{11} = \sqrt{a_{11}} \quad \Rightarrow \quad l_{11} = \sqrt{2}$$

$$\Rightarrow a_{12} = l_{11} l_{21} \quad \Rightarrow \quad l_{21} = \frac{a_{12}}{l_{11}} \quad \Rightarrow \quad l_{21} = -\sqrt{2}$$

$$\Rightarrow a_{22} = l_{21}^2 + l_{22}^2 \quad \Rightarrow \quad l_{22} = \sqrt{a_{22} - l_{21}^2} \quad \Rightarrow \quad l_{22} = \sqrt{5-2} = \sqrt{3}$$

$$\Rightarrow \boxed{A = \begin{bmatrix} \sqrt{2} & 0 \\ -\sqrt{2} & \sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix}}$$

2. (Lange Exercise 7.8) Suppose the matrix $A = \{a_{ij}\}$ is banded in the sense that $a_{ij} = 0$ when $|i - j| > d$. Prove that the Cholesky decomposition $B = \{b_{ij}\}$ also satisfies the band condition $b_{ij} = 0$ when $|i - j| > d$.

Small Bonus: what can we tell about sparse matrices in general?

Prove any banded $n \times n$ matrix A decomposes via Cholesky into banded $n \times n$ matrices

Proof by induction:

Base Case:

Let A be a banded $(d+1)\times(d+1)$ matrix s.t. $a_{ij}=0$ if $|i-j|>d$ i.e. at $a_{(d+1),1}$ and $a_{1,(d+1)}$

then Cholesky decomposition gives $A = BB^T$

where $B = \begin{bmatrix} b_{11} & \underline{0} \\ \underline{b} & B_{22} \end{bmatrix}$ and $\underline{b} = \left[\frac{a_{21}}{\sqrt{a_{11}}} \cdots \frac{a_{d1}}{\sqrt{a_{11}}} \;\; 0\right]^T$

thus $b_{ij}=0$ if $|i-j|>d$ i.e. at $b_{(d+1),1}$ and $b_{1,(d+1)}$ so B is banded

Induction Step

Assume any $n \times n$ banded matrix decomposes via Cholesky into banded matrices

Let A be a banded $(n+1)\times(n+1)$

Define A's structure as $A = \begin{bmatrix} a_{11} & \underline{a}^T \\ \underline{a} & A_{22} \end{bmatrix}$

Note: A_{22} is also banded, which can easily be verified by contradiction

Then A can be decomposed via cholesky to $A = BB^T$

where $B = \begin{bmatrix} b_{11} & \underline{0} \\ \underline{b} & B_{22} \end{bmatrix}$ and $\underline{b} = [\underbrace{\frac{a_{21}}{\sqrt{a_{11}}} \cdots \frac{a_{d1}}{\sqrt{a_{11}}}}_{d-1} \; \underbrace{0 \cdots 0}_{n-d}]^T$

and $B_{22}B_{22}^T = A_{22} - \underline{b}\underline{b}^T$

$A_{22} - \underline{b}\underline{b}^T$ is banded so it's Cholesky decomposition $B_{22}B_{22}^T$ is also banded by inductive assumption

Invoking induction, we have that any banded matrix decomposes into the product of banded matrices by Cholesky decomposition

3. (Lange Exercise 7.11) If $X = QR$ is the QR decomposition of X, where X has linearly independent columns, then show that the projection matrix

$$X(X^TX)^{-1}X^T = QQ^T.$$

In addition, show that $|\det(X)| = |\det(R)|$ when X is square and in general that $\det(X^TX) = [\det(R)]^2$.

Note: $X^TX = (QR)^TQR = R^T\underbrace{Q^TQ}_{I \text{ because orthonormal}}R = R^TR$

$$X\underbrace{(X^TX)^{-1}}_{=(R^TR)^{-1}}X^T = QQ^T$$

$\Rightarrow (R^TR)^{-1} = X^{-1}QQ^T(X^T)^{-1}$ X has independent columns so is invertible

$\Rightarrow (R^TR)^{-1} = (QR)^{-1}QQ^T(R^TQ^T)^{-1}$

$\Rightarrow (R^TR)^{-1} = R^{-1}\underbrace{Q^{-1}Q}_{=I}\underbrace{Q^T(Q^T)^{-1}}_{=I}(R^T)^{-1}$

$\Rightarrow (R^TR)^{-1} = R^{-1}(R^T)^{-1}$

$\Rightarrow (R^TR)^{-1} = (R^TR)^{-1}$ ■

$$|\det(X)| = |\det(QR)|$$

$$= |\underbrace{\det(Q)}_{=\pm 1}\det(R)| \quad \det(Q) = \pm 1 \text{ because it is orthogonal when } X \text{ is square}$$

$$= |(\pm 1)\det(R)|$$

$$= |\det(R)| \quad \text{when } X \text{ is square, so } \det(X) \text{ is defined}$$

$$|\det(X^T X)| = |\det(R^T R)| \quad \text{from above}$$

$$= |\underbrace{\det(R^T)}_{=\det(R)}\det(R)|$$

$$= (\det(R))^2 \quad \blacksquare$$

4. (Lange Exercise 8.4, modified) Show that the reflection matrix

$$A = \begin{pmatrix} cos(\theta) & sin(\theta) \\ sin(\theta) & -cos(\theta) \end{pmatrix}$$

is orthogonal and find its eigenvalues and eigenvectors.

show $A^T A = \alpha I$ for some scalar α

$$A^T A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) & \cos^2(\theta) + \sin^2(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \quad \text{so } A \text{ is orthogonal, } \alpha = 1$$

Find eigenvalues and eigenvectors

$$Ax = \lambda x \Rightarrow (A - \lambda I)x = 0 \Rightarrow |A - \lambda I| = 0$$

$$\left|\begin{bmatrix} \cos(\theta) - \lambda & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) - \lambda \end{bmatrix}\right| = (\cos(\theta) - \lambda)(-\cos(\theta) - \lambda) - \sin^2(\theta)$$

$$= -\cos^2(\theta) + \lambda^2 - \sin^2(\theta)$$

$$= -1 + \lambda^2 = 0$$

$$\Rightarrow \lambda^2 = 1$$

$$\Rightarrow \boxed{\lambda = \pm 1}$$

Find eigenvector associated with $\lambda = 1$

$$\begin{bmatrix} \cos(\theta)-1 & \sin(\theta) \\ \sin(\theta) & -\cos(\theta)-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (\cos(\theta)-1)x_1 + \sin(\theta)x_2 \\ \sin(\theta)x_1 - (\cos(\theta)+1)x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad x_1 = \frac{\sin(\theta)x_2}{1-\cos(\theta)} = \frac{\sqrt{1-\cos^2(\theta)}}{1-\cos(\theta)}x_2 = \frac{\sqrt{1+\cos(\theta)}\sqrt{1-\cos(\theta)}}{1-\cos(\theta)}x_2 = \sqrt{\frac{1+\cos(\theta)}{1-\cos(\theta)}}x_2$$

$$\Rightarrow \boxed{\begin{bmatrix} \sqrt{\frac{1+\cos(\theta)}{1-\cos(\theta)}} \\ 1 \end{bmatrix} \text{ is an eigen vector when } \lambda = 1}$$

Find eigenvector associated with $\lambda = -1$

$$\begin{bmatrix} \cos(\theta)+1 & \sin(\theta) \\ \sin(\theta) & -\cos(\theta)-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow (\cos(\theta)+1)x_1 + \sin(\theta)x_2 = 0$$

$$\Rightarrow x_1 = -\sqrt{\frac{1+\cos(\theta)}{1-\cos(\theta)}}x_2$$

$$\Rightarrow \boxed{\begin{bmatrix} -\sqrt{\frac{1+\cos(\theta)}{1-\cos(\theta)}} \\ 1 \end{bmatrix} \text{ is an eigen vector when } \lambda = -1}$$

5. (Lange Exercise 8.5, modified) Suppose λ is an eigenvalue of the orthogonal matrix O with corresponding eigenvector $\mathbf{v}$. Show that if $\mathbf{v}$ has real valued entries, then $\lambda = \pm 1$. **Hint**: use properties of the norms of orthogonal matrices.

$Ov = \lambda v$

$(Ov)^T Ov = (\lambda v)^T \lambda v$

$\Rightarrow v^T \underbrace{O^T O}_{=I} v = \lambda^2 v^T v$

$\Rightarrow v^T v = \lambda^2 v^T v$ $\quad$ $v^T v \neq 0$ because v is an eigenvector

$\Rightarrow \lambda^2 = 1$

$\Rightarrow \lambda = \pm 1$ $\quad$ That doesn't solve the realness of v so try another approach:

$Ov = \lambda v \Rightarrow \|Ov\|^2 = \|\lambda v\|^2$

$\Rightarrow \overline{(Ov)}^T (Ov) = |\lambda|^2 \|v\|^2$

$\Rightarrow \bar{v}^T \underbrace{\bar{O}^T O}_{=I} v = |\lambda|^2 \bar{v}^T v$

$\Rightarrow \bar{v}^T v = |\lambda|^2 \bar{v}^T v$ $\quad$ assuming v is real-valued

$\Rightarrow v^T v = |\lambda|^2 v^T v$

$\Rightarrow \lambda^2 = 1$

$\Rightarrow \boxed{\lambda = \pm 1}$

6. (Lange Exercise 9.3, modified) It can be shown that the matrix norm induced by the Euclidean (L_2) norm of matrix A is equal to the largest singular value of this matrix. Now, let A be an invertible $m \times m$ matrix with singular values $\sigma_1, \cdots, \sigma_m$. Recall that the L_2 condition number is defined as $\text{cond}_2(A) = ||A||_2||A^{-1}||_2$. Prove that

$$\text{cond}_2(A) = \frac{\max_i \sigma_i}{\min_i \sigma_i}.$$

Note: $A = V\Sigma U^T$ where U and V are orthonormal square $m\times m$ matrices and $\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix}$

$\Rightarrow$ $A^{-1} = (U^T)^{-1} \Sigma^{-1} V^{-1}$ which is the SVD of A^{-1} where $\Sigma = \begin{bmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_m} \end{bmatrix}$

$\Rightarrow$ $||A^{-1}||_2 = \max_i \frac{1}{\sigma_i} = \frac{1}{\min_i \sigma_i}$

Also: $||A||_2 = \max_i \sigma_i$

$\Rightarrow$ $\text{cond}_2(A) = ||A||_2 \, ||A^{-1}||_2 = (\max_i \sigma_i)\left(\frac{1}{\min_i \sigma_i}\right) = \boxed{\frac{\max_i \sigma_i}{\min_i \sigma_i}}$