496 Mathematics for Machine Learning and Inference

Problem Sheet for Tutorial 3

Problem 1
Let matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -2 & -3 & 1 \end{bmatrix}$$

perform QR decomposition on $\mathbf{A}$.

Problem 2
Prove the following:

- (1) A triangular matrix $\mathbf{A}^{n\times n}$ is normal (i.e. $\mathbf{A}^\intercal\mathbf{A} = \mathbf{A}\mathbf{A}^\intercal$) iff it is diagonal (just prove for an upper triangular).
- (2) If a triangular matrix $\mathbf{A}^{n\times n}$ is unitary and $a_{ii} > 0$, $i = 1, \dots, n$ then $\mathbf{A} = \mathbf{I}$.
- (3) Show that the inverse of an upper triangular matrix is upper triangular and that the product of two upper triangular matrices is upper triangular.
- (4) Based on (1, 2, 3) prove that the QR decomposition of a matrix $\mathbf{A}^{n\times n}$ (i.e. $\mathbf{A} = \mathbf{QR}$) with $r_{ii} > 0$, $\forall\, i = 1, \dots, n$ is unique.

Problem 3

- (1) Show that the absolute value of the determinant of a unitary matrix equals 1.
- (2) Prove that the determinant of a upper triangular square matrix is equal to the product of its diagonal elements.
- (3) Using (3) and (3), show that the absolute value of the determinant of an $n \times n$-matrix $\mathbf{A}$ with QR decomposition $\mathbf{A} = \mathbf{QR}$ equals to the product of diagonal elements of $\mathbf{R}$, i.e. $|\det\mathbf{A}| = \prod_{i=1}^{n} r_{ii}$

Solution

Problem 1

$A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$, then $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ \text{-}2 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \\ \text{-}3 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

Let $\mathbf{Q} = [\mathbf{w}_1 \ \ \mathbf{w}_2 \ \ \mathbf{w}_3]$, then since $||\mathbf{a}_1|| = \sqrt{6}$, $\mathbf{w}_1 = \mathbf{a}_1/||\mathbf{a}_1|| = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \text{-}\frac{2}{\sqrt{6}} \end{bmatrix}$.

First we compute $< \mathbf{w}_1, \mathbf{a}_2 >= \mathbf{w}_1^\intercal \mathbf{a}_2 = 9/\sqrt{6}$

$$\mathbf{w}_2 = \mathbf{a}_2 - < \mathbf{w}_1, \mathbf{a}_2 > \mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \\ \text{-}3 \end{bmatrix} - \frac{9}{6} \begin{bmatrix} 1 \\ 1 \\ \text{-}2 \end{bmatrix} = \begin{bmatrix} \text{-}\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

and $||\mathbf{w}_2|| = \sqrt{1/2}$. Then, $\mathbf{w}_2 = \frac{\mathbf{w}_2}{||\mathbf{w}_2||} = \begin{bmatrix} \text{-}\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$

$< \mathbf{a}_3, \mathbf{w}_1 >= -\frac{1}{\sqrt{6}}$ and $< \mathbf{a}_3, \mathbf{w}_2 >= \frac{1}{\sqrt{2}}$. Then,

$$\begin{aligned} \mathbf{w}_3 &= \mathbf{a}_3 - < \mathbf{a}_3, \mathbf{w}_1 > \mathbf{w}_1 - < \mathbf{a}_3, \mathbf{w}_2 > \mathbf{w}_2 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} 1 \\ 1 \\ \text{-}2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} \text{-}1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \end{aligned} \tag{1}$$

$$||\mathbf{w}_3|| = 2/\sqrt{3} \ \text{ hence } \ \mathbf{w}_3 = \frac{\mathbf{w}_3}{||\mathbf{w}_3||} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \tag{2}$$

Hence,

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \text{-}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \text{-}\frac{1}{\sqrt{3}} \end{bmatrix}$$

and

$$\mathbf{R} = \begin{bmatrix} < \mathbf{w}_1, \mathbf{a}_1 > & < \mathbf{w}_1, \mathbf{a}_2 > & < \mathbf{w}_1, \mathbf{a}_3 > \\ 0 & < \mathbf{w}_2, \mathbf{a}_2 > & < \mathbf{w}_2, \mathbf{a}_3 > \\ 0 & 0 & < \mathbf{w}_3, \mathbf{a}_3 > \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \frac{9}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$

Problem 2

(1) We will use induction.

- For $n = 1$ there is nothing to prove.
- Suppose that it is the case for all upper triangular matrices up to $n = k - 1$ with $k > 2$, i.e. $n = 2, 3, \ldots, k - 1$.
- For $n = k$, let us first express the upper triangular matrix in block form as:

$$\mathbf{A}_k = \left[\begin{array}{c|c} a_{1,1} & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & \tilde{\mathbf{A}} \end{array}\right] \tag{3}$$

where $a_{1,1}$ is a scalar, $\mathbf{z}$ is a $k - 1 \times 1$ vector, $\mathbf{0}$ is a $k - 1 \times 1$ vector with zeros and $\tilde{\mathbf{A}}$ is a $k - 1 \times k - 1$ upper triangular matrix.

$$\mathbf{A}_k\mathbf{A}_k^\mathsf{T} = \left[\begin{array}{c|c} a_{1,1} & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & \tilde{\mathbf{A}} \end{array}\right] \left[\begin{array}{c|c} a_{1,1} & \mathbf{0} \\ \hline \mathbf{z} & \tilde{\mathbf{A}}^\mathsf{T} \end{array}\right] = \left[\begin{array}{c|c} a_{1,1}^2 + \mathbf{z}^\mathsf{T}\mathbf{z} & \mathbf{z}^\mathsf{T}\tilde{\mathbf{A}}^\mathsf{T} \\ \hline \tilde{\mathbf{A}}\mathbf{z} & \tilde{\mathbf{A}}\tilde{\mathbf{A}}^\mathsf{T} \end{array}\right] \tag{4}$$

$$\mathbf{A}_k^\mathsf{T}\mathbf{A}_k = \left[\begin{array}{c|c} a_{1,1} & \mathbf{0} \\ \hline \mathbf{z} & \tilde{\mathbf{A}}^\mathsf{T} \end{array}\right] \left[\begin{array}{c|c} a_{1,1} & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & \tilde{\mathbf{A}} \end{array}\right] = \left[\begin{array}{c|c} a_{1,1}^2 & a_{1,1}\mathbf{z}^\mathsf{T} \\ \hline a_{1,1}\mathbf{z} & \mathbf{z}\mathbf{z}^\mathsf{T} + \tilde{\mathbf{A}}^\mathsf{T}\tilde{\mathbf{A}} \end{array}\right] \tag{5}$$

In order for $\mathbf{A}$ to be normal, $\mathbf{A}_k\mathbf{A}_k^\mathsf{T} = \mathbf{A}_k^\mathsf{T}\mathbf{A}_k$ must be true, thus:

$$\begin{cases} a_{1,1}^2 = a_{1,1}^2 + \mathbf{z}^\mathsf{T}\mathbf{z} \Rightarrow \mathbf{z}^\mathsf{T}\mathbf{z} = 0 \Rightarrow \mathbf{z} = \mathbf{0} \\ \mathbf{z}^\mathsf{T}\tilde{\mathbf{A}}^\mathsf{T} = \tilde{\mathbf{A}}\mathbf{z} = a_{1,1}\mathbf{z}^\mathsf{T} = a_{1,1}\mathbf{z} = \mathbf{0} \end{cases} \tag{6}$$

And from the lower-right blocks we get:

$$\tilde{\mathbf{A}}\tilde{\mathbf{A}}^\mathsf{T} = \mathbf{z}\mathbf{z}^\mathsf{T} + \tilde{\mathbf{A}}^\mathsf{T}\tilde{\mathbf{A}} \Rightarrow \tilde{\mathbf{A}}\tilde{\mathbf{A}}^\mathsf{T} = \tilde{\mathbf{A}}^\mathsf{T}\tilde{\mathbf{A}} \tag{7}$$

which means that $\tilde{\mathbf{A}}$ is a $k-1\times k-1$ upper triangular normal matrix, hence diagonal (as assumed in the previous step). Consequently, $\mathbf{A}_k$ is diagonal.

(2) We denote $\mathbf{A}_k = \mathbf{A}^{k\times k}$. Let's solve the problem using the technique of induction on k. Therefore,

1. The base case ($k = 1$) is simple, as all scalars are trivially "upper triangular".

2. For $k=2$, $\mathbf{A}_2$ is upper trinagular, $\mathbf{A}_2\mathbf{A}_2^\mathsf{T}=\mathbf{I}_2$ and $a_{ii}>0$, $i=1,2$. We have that

$$\mathbf{A}_2\mathbf{A}_2^\mathsf{T}=\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}\begin{bmatrix} a_{11} & 0 \\ a_{12} & a_{22} \end{bmatrix}=\begin{bmatrix} a_{11}^2+a_{12}^2 & a_{12}a_{22} \\ a_{12}a_{22} & a_{22}^2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\Rightarrow$$
$$\begin{cases} a_{22}=1 \\ a_{12}=0 \\ a_{11}=1 \end{cases}\Rightarrow \mathbf{A}_2=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=\mathbf{I}_2$$

3. For $k=n$, we assume that if $\mathbf{A}_n$ is unitary and upper triangular matrices with positive diagonal, then it equals to the identity matrix, i.e. $\mathbf{A}_n\mathbf{A}_n^\mathsf{T}=\mathbf{I}_n$, $\mathbf{A}_n=\mathbf{I}$

4. For $k=n+1$, let $\mathbf{A}_{n+1}$ be a unitary and upper triangular matrix with positive diagonal. In block form, then, we have

$$\mathbf{A}_{n+1}=\left[\begin{array}{c|c} \mathbf{A}_n & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & a \end{array}\right]$$

where $\mathbf{A}_n$ is a unitary and upper triangular matrices with positive diagonal, $\mathbf{0}$ is the $1\times n$ vector of 0s, $\mathbf{z}^\mathsf{T}$ is a $n\times 1$ vector and a is some nonzero scalar. The transpose of $\mathbf{A}_{n+1}$ is

$$\mathbf{A}_{n+1}^\mathsf{T}=\left[\begin{array}{c|c} \mathbf{A}_n^\mathsf{T} & \mathbf{0}^\mathsf{T} \\ \hline \mathbf{z} & a \end{array}\right]$$

Since, $\mathbf{A}_{n+1}$ is unitary we have

$$\begin{aligned}\mathbf{A}_{n+1}\mathbf{A}_{n+1}^\mathsf{T}&=\left[\begin{array}{c|c} \mathbf{A}_n & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & a \end{array}\right]\left[\begin{array}{c|c} \mathbf{A}_n^\mathsf{T} & \mathbf{0}^\mathsf{T} \\ \hline \mathbf{z} & a \end{array}\right]= \\ &=\left[\begin{array}{c|c} \mathbf{A}_n\mathbf{A}_n^\mathsf{T}+\mathbf{z}^\mathsf{T}\mathbf{z} & a\mathbf{z}^\mathsf{T} \\ \hline a\mathbf{z} & a^2 \end{array}\right]=\left[\begin{array}{c|c} \mathbf{I}_n & \mathbf{0}^\mathsf{T} \\ \hline \mathbf{0} & 1 \end{array}\right]\Rightarrow \\ &\begin{cases} a=1,\ \text{since}\, a>0 \\ \mathbf{z}=\mathbf{0} \\ \mathbf{A}_n=\mathbf{I}_n,\ \text{by step 3} \end{cases}\Rightarrow \mathbf{A}_{n+1}=\left[\begin{array}{c|c} \mathbf{I}_n & \mathbf{0}^\mathsf{T} \\ \hline \mathbf{0} & a \end{array}\right]=\mathbf{I}_{n+1}\end{aligned}$$

(3) We will use induction.

- For $n=1$ there is nothing to prove.
- Suppose that all upper triangular matrices $\mathbf{A}$ up to $n=k-1$ with $k>2$ have uper triangular inverse matrices.
- For $n=k$, let us first express the upper triangular matrix in block form as:

$$\mathbf{A}_k=\left[\begin{array}{c|c} a_{1,1} & \mathbf{z}^\mathsf{T} \\ \hline \mathbf{0} & \tilde{\mathbf{A}} \end{array}\right] \tag{8}$$

where $a_{1,1}$ is a scalar, $\mathbf{z}$ is a $k-1\times 1$ vector, $\mathbf{0}$ is a $k-1\times 1$ vector with zeros and $\tilde{\mathbf{A}}$ is a $k-1\times k-1$ upper triangular matrix. Following the same block form, the inverse matrix $\mathbf{A}^{-1}$ can be expressed as:

$$\mathbf{A}_k^{-1} = \left[\begin{array}{c|c} a'_{1,1} & \mathbf{z}'^{\mathsf{T}} \\ \hline \mathbf{x} & \tilde{\mathbf{A}}' \end{array}\right] \tag{9}$$

where $a'_{1,1}$ is a scalar, $\mathbf{z}'$ and $\mathbf{b}$ are $k-1\times 1$ vectors and $\tilde{\mathbf{A}}'$ is a $k-1\times k-1$ matrix. Similarly, let us write a $k\times k$ identity matrix in block form as:

$$\mathbf{I}_k = \left[\begin{array}{c|c} 1 & \mathbf{0}^{\mathsf{T}} \\ \hline \mathbf{0} & \mathbf{I_{k-1}} \end{array}\right] \tag{10}$$

where $\mathbf{0}$ is a $k-1\times 1$ vector with zeros and $\mathbf{I_{k-1}}$ is a $k-1\times k-1$ identity matrix. Given that $\mathbf{A}_k^{-1}$ is the inverse of $\mathbf{A}_k$, we have:

$$\begin{aligned}
\mathbf{A}_k^{-1}\mathbf{A}_k = \mathbf{I_k} \quad &\Rightarrow \left[\begin{array}{c|c} a'_{1,1} & \mathbf{z}'^{\mathsf{T}} \\ \hline \mathbf{x} & \tilde{\mathbf{A}}' \end{array}\right] \left[\begin{array}{c|c} a_{1,1} & \mathbf{z}^{\mathsf{T}} \\ \hline \mathbf{0} & \tilde{\mathbf{A}} \end{array}\right] = \left[\begin{array}{c|c} 1 & \mathbf{0}^{\mathsf{T}} \\ \hline \mathbf{0} & \mathbf{I_{k-1}} \end{array}\right] \\
&\Rightarrow \left[\begin{array}{c|c} a'_{1,1}a_{1,1} & a'_{1,1}\mathbf{z}^{\mathsf{T}} + \mathbf{z}'^{\mathsf{T}}\tilde{\mathbf{A}} \\ \hline a_{1,1}\mathbf{x} & \mathbf{x}\mathbf{z}^{\mathsf{T}} + \tilde{\mathbf{A}}'\tilde{\mathbf{A}} \end{array}\right] = \left[\begin{array}{c|c} 1 & \mathbf{0}^{\mathsf{T}} \\ \hline \mathbf{0} & \mathbf{I_{k-1}} \end{array}\right] \\
&\Rightarrow \begin{cases} a'_{1,1}a_{1,1} = 1 \Rightarrow a'_{1,1} \text{ and } a_{1,1} \text{ are non-zero} \\ a_{1,1}\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0} \\ \mathbf{x}\mathbf{z}^{\mathsf{T}} + \tilde{\mathbf{A}}'\tilde{\mathbf{A}} = \mathbf{I_{k-1}} \Rightarrow \tilde{\mathbf{A}}'\tilde{\mathbf{A}} = \mathbf{I_{k-1}} \text{ previous step} \end{cases}
\end{aligned} \tag{11}$$

Thus, by inductive hypothesis, we have that

$$\mathbf{A}_k^{-1} = \left[\begin{array}{c|c} a'_{1,1} & \mathbf{z}'^{\mathsf{T}} \\ \hline \mathbf{0} & \tilde{\mathbf{A}}' \end{array}\right] \tag{12}$$

is upper triangular.

(4) From the properties of an upper triangular matrix $\mathbf{A}$ of size $n\times n$, an element $a_{i,j}$ is non-zero only if $i \leq j$, $\forall i = 1,\dots,n$, $\forall j = 1,\dots,n$. Now, let $\mathbf{A}$ and $\mathbf{B}$ be two upper triangular matrices of size $n\times n$. By definition, their product $\mathbf{C} = \mathbf{AB}$ can be denoted as

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} \tag{13}$$

For a pair of i, j values, this sum will be non-zero only if there is some k with $i \leq k \leq j$ (if $i < k < j$, $\forall k = 1, \ldots, n$, then at least one of $a_{i,k}$ and $b_{k,j}$ will be zero). Consequently, $c_{i,j}$ is non-zero only if $i \leq j$, which means that $\mathbf{C}$ is upper triangular.

(**e**) Let's assume that matrix $\mathbf{A}$ has two QR decompositions $\mathbf{Q}_1\mathbf{R}_1$, $\mathbf{Q}_1\mathbf{Q}_1^{\mathsf{T}} = \mathbf{I}$ and $\mathbf{Q}_2\mathbf{R}_2$, $\mathbf{Q}_2\mathbf{Q}_2^{\mathsf{T}} = \mathbf{I}$. Then $\mathbf{A} = \mathbf{Q}_1\mathbf{R}_1 = \mathbf{Q}_2\mathbf{R}_2$ and $\mathbf{Q}_1\mathbf{R}_1 = \mathbf{Q}_2\mathbf{R}_2 \Rightarrow \mathbf{R}_2^{-1}\mathbf{R}_1 = \mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2$.
We can readily see that

$$\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2(\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2)^{\mathsf{T}} = \mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2\mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1 = \mathbf{I}$$
$$(\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2)^{\mathsf{T}}\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2 = \mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2 = \mathbf{I}$$

Thus, the matrix $\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2$ is normal and unitary and so it is the matrix $\mathbf{R}_2^{-1}\mathbf{R}_1$. According to (**2**) and (**??**), the inverse of a triangular matrix is also a triangular matrix and the product of triangular matrices is a triangular matrix as well. Therefore, $\mathbf{R}_2^{-1}\mathbf{R}_1$ is unitary and diagonal with positive diagonal (because the main diagonal entries of both $\mathbf{R}_2^{-1}$ and $\mathbf{R}_1$ are positive) which based on (**2**) implies that it equals to the identity matrix. Hence, $\mathbf{R}_2^{-1}\mathbf{R}_1 = \mathbf{I} \Rightarrow \mathbf{R}_1 = \mathbf{R}_2$ and $\mathbf{Q}_1 = \mathbf{Q}_2$

Problem 3
(1) $\mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \mathbf{I} \Rightarrow \det^2\mathbf{Q} = 1$. So, $|\det\mathbf{Q}| = 1$
(2) $\forall r \in [1\,..\,n]\, \det\mathbf{A} = \sum_{k=1}^{n} a_{rk}\,(-1)^{r+k}\det_{rk}\mathbf{A}$ where $\det_{rk}\mathbf{A}$ is the $n-1$ determinant obtained from $\mathbf{A}$ by deleting row r and column k.
We proceed by induction on k, the number of rows of $\mathbf{A}$.

- For $k = 1$, the determinant is a_{11}, which is clearly also the diagonal element. This forms the basis for the induction.
- Let us assume that for $k = n$, $\mathbf{A}_n$ is upper triagonal matrix and its determinant is $\det\mathbf{A}_n = \prod_{i=1}^{n} a_{ii}$
- For $k = n+1$, $\mathbf{A}_{k+1}$ is upper triagonal matrix and its determinant is $\det\mathbf{A}_{n+1} = \sum_{k=1}^{n+1} a_{r\,k}\,(-1)^{r+k}\det_{r\,k}\mathbf{A}$, $\forall r \in [1\,..\,n+1]$. We notice that for $r = n+1$, $a_{rk} \neq 0$ only for $k = n+1$. Therefore,

$$\det\mathbf{A}_{n+1} = a_{n+1,n+1}\,(-1)^{n+1+n+1}\det_{n+1\,n+1}\mathbf{A},$$

  By definition, $\det_{n+1\,n+1}\mathbf{A} = \det\mathbf{A}_n$.

  But $\det\mathbf{A}_n = \prod_{i=1}^{n} a_{ii}$. So,

$$\det \mathbf{A}_{n+1} = a_{n+1,n+1} \prod_{i=1}^{n} a_{ii} = \prod_{i=1}^{n+1} a_{ii}$$

and the result follows by induction.

(3) $|\det \mathbf{A}| = |\det \mathbf{Q} \det \mathbf{R}| = |\det \mathbf{Q}|\,|\det \mathbf{R}| = |\det \mathbf{R}| = \displaystyle\prod_{i=1}^{n} r_{ii}$