Exercise 2 - derivations

Ditte Grønborg Blom, Viktor Stenby Johansson

23. november 2022

Lemma 1
For an orthogonal matrix $\mathbf{Q}$ and a standard Gaussian random matrix $\mathbf{\Omega}$, then the product $\mathbf{Q\Omega}$ is also a standard Gaussian random matrix.

Lemma 2
Any submatrix of a standard Gaussian random matrix is a standard Gaussian random matrix.

For matrices $\mathbf{A}$, $\mathbf{X}$ and $\mathbf{Y}$, we have:

$$\mathbf{A} \ \in \mathbb{R}^{n\times m}, \quad \mathbf{X} \in \mathbb{R}^{m\times r}, \quad \mathbf{Y} \in \mathbb{R}^{n\times r}$$

Using the singular value decomposition, $\mathbf{A}$ can be decomposed into orthonormal matrix $\mathbf{U} \in \mathbb{R}^{n\times n}$, $\mathbf{\Sigma} \in \mathbb{R}^{n\times m}$ and orthonormal matrix $\mathbf{V} \in \mathbb{R}^{n\times n}$ as follows:

$$\mathbf{A} = \mathbf{U\Sigma V}^T$$

Partitioning $\mathbf{U}$ and $\mathbf{V}$ as $\mathbf{U} = \begin{bmatrix} \mathbf{U}_r & \mathbf{U}_\perp \end{bmatrix}$, where $\mathbf{U}_r \ \in \mathbb{R}^{n\times r}$ and similarly $\mathbf{V} = \begin{bmatrix} \mathbf{V}_r & \mathbf{V}_\perp \end{bmatrix}$, where $\mathbf{V}_r \in \mathbb{R}^{m\times r}$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{U}_r & \mathbf{U}_\perp \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_r & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(m-r)} \end{bmatrix} \begin{bmatrix} \mathbf{V}_r^T \\ \mathbf{V}_\perp^T \end{bmatrix} = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^T$$

where we used that $\mathbf{A}$ is of rank r and therefore only has r singular values.

Proposition 1
If $\mathbf{X} \in \mathbb{R}^{m\times r}$ is a standard Gaussian random matrix, then $\mathbf{V_r}^T\mathbf{X} \in \mathbb{R}^{r\times r}$ is a standard Gaussian random matrix.

Proof
$\mathbf{V}$ and $\mathbf{V}^T$ are both orthonormal matrices. Therefore $\mathbf{V}^T\mathbf{X}$ is a standard Gaussian random matrix. If we use the partitioning of $\mathbf{V}$ as introduced earlier, we have:

$$\mathbf{V}^T\mathbf{X} = \begin{bmatrix} \mathbf{V}_r^T \\ \mathbf{V}_\perp^T \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathbf{V}_r^T\mathbf{X} \\ \mathbf{V}_\perp^T\mathbf{X} \end{bmatrix} \tag{1}$$

Recognizing $\mathbf{V}_r^T\mathbf{X}$ as a submatrix of a standard Gaussian random matrix and using Lemma 2 means that $\mathbf{V}_r^T\mathbf{X}$ is a standard Gaussian random matrix. □

Proposition 2
If $\mathbf{Y} \in \mathbb{R}^{n \times r}$ is a standard Gaussian random matrix, then $\mathbf{Y}^T\mathbf{U}_r$ is a standard Gaussian random matrix.

Proof
$\mathbf{U}$ and $\mathbf{U}^T$ are both orthonormal matrices. Lemma 1 states that $\mathbf{U}^T\mathbf{Y}$ is a standard Gaussian random matrix, meaning that $\mathbf{Y}^T\mathbf{U} = \left(\mathbf{U}^T\mathbf{Y}\right)^T$ is a standard Gaussian random matrix, because the transpose of a standard Gaussian random matrix is also standard Gaussian. Writing

$$\mathbf{Y}^T\mathbf{U} = \mathbf{Y}^T \begin{bmatrix} \mathbf{U}_r & \mathbf{U}_\perp \end{bmatrix} = \begin{bmatrix} \mathbf{Y}^T\mathbf{U}_r & \mathbf{Y}^T\mathbf{U}_\perp \end{bmatrix} \tag{2}$$

allows us to recognize $\mathbf{Y}^T\mathbf{U}_r$ as a submatrix of a standard Gaussian random matrix, which according to Lemma 2 states that $\mathbf{Y}^T\mathbf{U}_r$ itself then is a standard Gaussian random matrix. $\square$

Proposition 3
If $\mathbf{A}$ has exactly rank r, then $\texttt{SKETCHING}(\mathbf{A}, r, 0)$ returns an exact low-rank factorization with probability 1.

Using the `SKETCHING` expression for $\mathbf{A}$ yields:

$$\begin{aligned}
\texttt{SKETCHING}(\mathbf{A}, r, 0) &= \mathbf{B}\mathbf{C}^T && (3)\\
&= \mathbf{A}\mathbf{X}\mathbf{R}^\dagger(\mathbf{A}^T\mathbf{Y}\mathbf{Q})^T && (4)\\
&= \mathbf{A}\mathbf{X}\mathbf{R}^\dagger\mathbf{Q}^T\mathbf{Y}^T\mathbf{A} && (5)\\
&= \mathbf{A}\mathbf{X}(\mathbf{Q}\mathbf{R})^\dagger\mathbf{Y}^T\mathbf{A} && (6)\\
&= \mathbf{A}\mathbf{X}\left(\mathbf{Y}^T\mathbf{A}\mathbf{X}\right)^\dagger\mathbf{Y}^T\mathbf{A} && (7)\\
&= \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T\mathbf{X}\left(\mathbf{Y}^T\mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T\mathbf{X}\right)^\dagger\mathbf{Y}^T\mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T && (8)\\
&= \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T\mathbf{X}\left(\mathbf{V}_r^T\mathbf{X}\right)^\dagger(\mathbf{\Sigma}_r)^\dagger\left(\mathbf{Y}^T\mathbf{U}_r\right)^\dagger\mathbf{Y}^T\mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T && (9)\\
&= \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T && (10)
\end{aligned}$$

where a bunch of stuff was used, which will be explained later.