This handout includes space for every question that requires a written response. Please feel free to use it to handwrite your solutions (legibly, please). If you choose to typeset your solutions, the `README.md` for this assignment includes instructions to regenerate this handout with your typeset LaTeX solutions.

4.a

Let $\mathbf{A} \in \mathbb{R}^{n\times d}$ with SVD $\mathbf{A} = \mathbf{UDV}^T$, where $\mathbf{U} \in \mathbb{R}^{n\times r}$, $\mathbf{D} \in \mathbb{R}^{r\times r}$, and $\mathbf{V}^{d\times r}$. Let $u_{ij} = (\mathbf{U})_{ij}$, $v_{ij} = (\mathbf{V})_{ij}$, and $\sigma_i = \mathbf{D}_{ii}$.

Note that

$$(\mathbf{UD})_{ij} = \sum_{k=1}^{r} u_{ik}\delta_{kj}\sigma_j = u_{ij}\sigma_j$$

where δ_{kj} is the Kronecker delta function. From this, we get

$$(\mathbf{UDV}^T)_{ij} = \sum_{k=1}^{r} (\mathbf{UD})_{ik}(\mathbf{V}^T)_{kj} = \sum_{k=1}^{r} \sigma_k u_{ik} v_{jk}.$$

Let $\mathbf{u}_i$ and $\mathbf{v}_i$ be the ith columns of $\mathbf{U}$ and $\mathbf{V}$, respectively. Then kl-th entry of $\mathbf{u}_i\mathbf{v}_i^T$ is $u_{ki}v_{li}$. Hence

$$\left(\sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\right)_{kl} = \sum_{i=1}^{r} \sigma_i (\mathbf{u}_i\mathbf{v}_i^T)_{kl} = \sum_{i=1}^{r} \sigma_i u_{ki} v_{li},$$

which is precisely the expression found above for $(\mathbf{UDV}^T)_{kl}$.

4.b

By part (a), we know

$$\mathbf{A} = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^T.$$

Using this expression for A combined with the orthonormality of the columns of $\mathbf{V}$, we see

$$\mathbf{A}\mathbf{v_i} = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^T \mathbf{v}_i = \sigma_i \mathbf{u}_i.$$

Thus scaling by $1/\sigma_i$ gives the desired result.

4.c

In part (a), we have already seen that row s of $\mathbf{A}$, denoted $\mathbf{a}_s$, can be written as

$$\mathbf{a}_s = \begin{bmatrix} \sum_{i=1}^{r} \sigma_i u_{si} v_{1i} & \sum_{i=1}^{r} \sigma_i u_{si} v_{2i} & \cdots & \sum_{i=1}^{r} \sigma_i u_{si} v_{di} \end{bmatrix} = \sum_{i=1}^{r} \sigma_i u_{s_i} \mathbf{v}_i^T,$$

so by definition of $\mathbf{A}_k$, we set the all but the k largest singular values to zero and get an expression for row s of $\mathbf{A}_k$, denoted by $\mathbf{a}_{ks}$,

$$\mathbf{a}_{ks} = \sum_{i=1}^{k} \sigma_i u_{s_i} \mathbf{v}_i^T.$$

Using the hint and the fact that the $\mathbf{v}_i$ form an orthonormal basis for V_k, the projections onto $\mathbf{v}_i$ are simply

$$(\mathbf{a}_s \cdot \mathbf{v}_i)\mathbf{v}_i = \sigma_i u_{si} \mathbf{v}_i,$$

so we get

$$\mathbf{a}_{ks} = \sum_{i=1}^{k} \sigma_i u_{s_i} \mathbf{v}_i^T = \sum_{i=1}^{k} (\mathbf{a}_s \cdot \mathbf{v}_i)\mathbf{v}_i^T$$

as desired.

4.d

Since the square root is a strictly increasing function on $[0, \infty)$, we can equivalently minimize the square of the Frobenius norm.

Suppose $\mathbf{B}$ is a rank k matrix in $\mathbb{R}^{n \times d}$ and let $\mathbf{a}_s$ and $\mathbf{b}_s$ be the sth rows of $\mathbf{A}$ and $\mathbf{B}$, respectively. Let a_{ij} and b_{ij} be the ij-th entries of $\mathbf{A}$ and $\mathbf{B}$, respectively. Then the square of the Frobenius norm of $\mathbf{A} - \mathbf{B}$ is

$$\|\mathbf{A} - \mathbf{B}\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{d} (a_{ij} - b_{ij})^2 = \sum_{i=1}^{n} \|\mathbf{a}_i - \mathbf{b}_i\|_2^2.$$

Let B be the rowspace of $\mathbf{B}$. The projection of $\mathbf{a}_i$ onto B minimizes the distance between $\mathbf{a}_i$ and any vector in B, so if we want to find a rank k matrix $\mathbf{B}$ minimizing $\|\mathbf{A} - \mathbf{B}\|_F^2$, we can assume $\mathbf{b}_i = \mathrm{proj}_B(\mathbf{a}_i)$. However, this sum is minimized precisely when $B = V_k$ by properties of the SVD of $\mathbf{A}$. By part (c), we have $\mathrm{proj}_{V_k}(\mathbf{a}_i) = \mathbf{a}_{ki}$, the ith row of $\mathbf{A}_k$, so $\|\mathbf{A} - \mathbf{B}\|_F^2$ is minimized when $\mathbf{B} = \mathbf{A}_k$. (Note it is not unique if V_k does not contain the entire subspace generated by all singular vectors corresponding to the smallest singular value σ_k.)