Exercise 1 (SVD and the reconstruction problem)

Let X be a $n \times p$ matrix. Show that a solution to the following optimization problem

$$\min_{u \in \mathbb{R}^n, v \in \mathbb{R}^p} J(u,v) \qquad \text{with} \qquad J(u,v) \;=\; \|X - uv^\top\|_F^2$$

with $\|v^\star\| = 1$, is given by

$$v^\star \qquad \text{and} \qquad u^\star = X\, v^\star,$$

where $v^\star$ is the normalized eigen vector associated with λ the largest eigen value of $X^\top X$. Furthermore, $\|u^\star\| = \sqrt{\lambda}$.

Solution.

first note that $J(u,v) \;=\; \|X - uv^\top\|_F^2$ can be written:

$$\begin{aligned}
\sum_i^n \sum_j^p (x_{ij} - u_i v_j)^2 &= \sum_i^n \sum_j^p x_{ij}^2 - 2 \sum_i^n \sum_j^p x_{ij} u_i v_j + \sum_i^n \sum_j^p (u_i v_j)^2 \\
&= \sum_i^n \sum_j^p x_{ij}^2 - 2 \sum_i^n \left(\sum_j^p x_{ij} v_j \right) u_i + \sum_i^n u_i^2 \; \sum_j^p v_j^2 \\
&= \sum_i^n \sum_j^p x_{ij}^2 - 2 (Xv)^\top u + \|u\|^2 \; \|v\|^2
\end{aligned}$$

so that

$$\min_{u,v} \underbrace{\|X - uv^\top\|_F^2}_{J(u,v)} \quad \text{admit the same solution as} \quad \min_{u,v} \underbrace{-2(Xv)^\top u + \|u\|^2 \; \|v\|^2}_{\mathcal{J}(u,v)}$$

Minimize the cost amount at finding u et v canceling the gradient, that is:

$$\min_{u,v} J(u,v) \;=\; \|X - uv^\top\|_F^2 \qquad \Leftrightarrow \qquad \begin{cases} \nabla_u \mathcal{J}(u) = 0 \\ \nabla_v \mathcal{J}(v) = 0 \end{cases}$$

The gradients of $J(u,v)$ are:

$$\begin{cases}
\dfrac{\partial J(u,v)}{\partial u_i} &= -2 \displaystyle\sum_j^p x_{ij} v_j + 2 u_i \; \sum_j^p v_j^2 \\
\dfrac{\partial J(u,v)}{\partial v_j} &= -2 \displaystyle\sum_i^n x_{ij} u_i + 2 v_j \; \sum_i^n u_i^2
\end{cases}$$

$$\begin{cases} \nabla_u J(u) = -2Xv + 2\|v\|^2 u \\ \nabla_v J(v) = -2X^\top u + 2\|u\|^2 v \end{cases}$$

Optimality conditions are:

$$\begin{cases} \nabla_u \mathcal{J}(u) = 0 & \Leftrightarrow \quad -Xv + \|v\|^2 u = 0 \\ \nabla_v \mathcal{J}(v) = 0 & \Leftrightarrow \quad -X^\top u + \|u\|^2 v = 0 \end{cases}$$

so that

$$\begin{cases} Xv = \|v\|^2 u \\ X^\top u = \|u\|^2 v \end{cases} \quad \Rightarrow X^\top X v = \|v\|^2 X^\top u = \underbrace{\|v\|^2 \|u\|^2}_{\lambda} \; v$$

The solution is given by an eigen vector v of matrix $X^\top X$.

At the optimum

$$\mathcal{J}(u,v) \;=\; -2(Xv)^\top u + \|u\|^2 \; \|v\|^2 \qquad \text{and} \qquad Xv = \|v\|^2 u$$

so that

$$\begin{aligned}
\mathcal{J}(u,v) &= -2\|v\|^2 u^\top u + \|u\|^2 \; \|v\|^2 \\
&= -2\|v\|^2 \|u\|^2 + \|u\|^2 \; \|v\|^2 \\
&= -\|u\|^2 \; \|v\|^2 \qquad = -\lambda
\end{aligned}$$

The cost is the eigen value so that the minimum is given by the largest one.

□

Exercise 2 (Iterated SVD)

The weighted SVD of order q problem is

$$\min_{U \in \mathbb{R}^{n \times q}, V \in \mathbb{R}^{p \times q}} \|X - UV^\top\|_W^2$$

where

$$\|M\|_W^2 = \sum_{i,j} W_{i,j} M_{i,j}^2. \tag{1}$$

We assume here that W is mask that is $W_{i,j} \in 0, 1$.

Let $\odot$ be the component-wise multiplication, the iterated SVD algorithm is given by

- Start by an initial factorization $U_0 V_0^\top$.
- Iterate T time:
  - Compute the completed matrix $R_t = W \odot X + (1 - W) \odot (U_t V_t^\top)$ where 1 is a matrix filled with 1 everywhere.
  - Use the SVD to obtain a rank q factorization of R_t by $U_{t+1} V_{t+1}^\top$
- Use the last factorization $U_T V_T^\top$.

We let $\mathcal{L}(U, V) = \|X - UV^\top\|_W^2$ and define

$$\mathcal{L}^+(U, V, U', V) = \|W \odot X + (1 - W) \odot U'V'^\top - UV^\top\|^2$$

1. Verify that $\mathcal{L}(U, V) \leq \mathcal{L}^+(U, V, U', V')$ and $\mathcal{L}(U, V) = \mathcal{L}^+(U, V, U, V)$

Solution.

A simple computation yields

$$\begin{aligned}\mathcal{L}^+(U, V, U', V') &= \|W \odot X + (1 - W) \odot U'V'^\top - UV^\top\|^2 \\ &= \|X - UV^\top\|_W^2 + \|U'V'^\top - UV^\top\|_{1-W}^2 \qquad \geq \|X - UV^\top\|_W^2\end{aligned}$$

Note that when $U' = U$ and $V' = V$ $\|U'V'^\top - UV^\top\|_{1-W}^2 = 0$, hence the result. □

2. Prove that $\mathcal{L}(U_{t+1}, V_{t+1}) \leq \mathcal{L}(U_t, V_t)$. You may use the fact that the SVD can be used to minimize the problem

$$\min_{U \in \mathbb{R}^{n \times q}, V \in \mathbb{R}^{p \times q}} \|X - UV^\top\|^2.$$

Solution.

By construction,

$$\begin{aligned}\mathcal{L}(U_{t+1}, V_{t+1}) &\leq \mathcal{L}(U_{t+1}, U_{t+1}, U_t, V_t) \\ &= \min_{U \in \mathbb{R}^{n \times q}, V \in \mathbb{R}^{p \times q}} \|W \odot X + (1 - W) \odot U_t V_t^\top - UV^\top\|^2 \\ &= \min_{U \in \mathbb{R}^{n \times q}, V \in \mathbb{R}^{p \times q}} \mathcal{L}^+(U, V, U_t, V_t) \\ &\leq \mathcal{L}^+(U_t, V_t, U_t, V_t) = \mathcal{L}(U_t, V_t)\end{aligned}$$

□

Exercise 3 (Alternated least square)

The penalized weighted SVD of order q problem is

$$\min_{U \in \mathbb{R}^{n \times q}, V} \|X - UV^\top\|_W^2 + \lambda \|U\|^2 + \lambda \|V\|^2$$

with $\lambda \geq 0$ where

$$\|M\|_W^2 = \sum_{i,j} W_{i,j} M_{i,j}^2 \tag{2}$$

with $W_{i,j} \geq 0$.

To solve it using the penalized Alternating Least Square(ALS) the following sequence has to be implemented: initialize U and V with the SVD on the full matrix

loop

compute U that $\min_U \|X - UV^\top\|_W^2 + \lambda \|U\|^2$ with a fix V

compute V that $\min_V \|X - UV^\top\|_W^2 + \lambda \|V\|^2$ with a fix U

1. Show that
$$\min_U \|X - UV^\top\|_W^2 + \lambda\|U\|^2 \text{ with a fix } V$$
boils down to solving n least square problems of the form
$$\min_{u_i \in \mathbb{R}^q} \|x_i - Vu_i\|_{D_i}^2 + \lambda\|u_i\|^2, \qquad \text{for } i = 1, \dots, n$$
with $x_i = X_{i,\cdot}$ a $\mathbb{R}^p$ vectors, $D_i = diag(W_{i,\cdot})$ the diagonal matrix with vector $W_{i,j}, j = 1, \dots, p$ on the diagonal, $\|x\|_D^2 = \sum_j D_j x_j^2$, and whose solution is
$$u_i = (V^\top D_i V + \lambda I)^{-1} V^\top D_i x_i,$$
where I is the identity matrix of dimension q.

Solution.

The problem to solve is
$$\min_U \|X - VU^\top\|_W^2 + \lambda\|U\|^2$$
that is
$$\min_{U \in \mathbb{R}^{n\times q}} \sum_{i=1}^n \sum_{j=1}^p W_{ij}\|X_{ij} - \sum_{k=1}^q V_{jk}U_{ik}\| + \lambda \sum_{i=1}^n \sum_{k=1}^q \|U_{ik}\|^2$$
or
$$\min_{U \in \mathbb{R}^{n\times q}} \sum_{i=1}^n \|X_{i\cdot} - VU_{i\cdot}^\top\|_{diag(W_i)} + \lambda \sum_{i=1}^n \|U_{i\cdot}\|^2$$
that can be recasted into n independent problems
$$\min_{u_i \in \mathbb{R}^q} \|x_i - Vu_i\|_{D_i}^2 + \lambda\|u_i\|^2$$
for $i = 1, \dots, n$. This is a classical (regularized) least square problem whose solution is given by
$$u_i = (V^\top D_i V + \lambda I)^{-1} V^\top D_i x_i,$$

□

2. Show that
$$\min_V \|X - UV^\top\|_W^2 + \lambda\|V\|^2 \text{ with a fix } U$$
boils down at solving p least square problems of the form
$$\min_{v_j \in \mathbb{R}^q} \|x_j - Uv_j\|_{D_j}^2 + \lambda\|v_i\|^2, \qquad \text{for } j = 1, \dots, p$$
with $x_j = X_{\cdot,j}$ a $\mathbb{R}^n$ vectors and $D_j = diag(W_{\cdot,j})$ the diagonal matrix with vector $W(i,j), i = 1, \dots, n$ and whose solution is
$$v_j = (U^\top D_j U + \lambda I)^{-1} D^\top D_j x_j,$$
where I is the identity matrix of dimension q.

Solution.

The problem to solve is now
$$\min_U \|X - VU^\top\|_W^2 + \lambda\|V\|^2$$
that is
$$\min_{V \in \mathbb{R}^{q\times p}} \sum_{i=1}^n \sum_{j=1}^p W_{ij}\|X_{ij} - \sum_{k=1}^q V_{jk}U_{ik}\| + \lambda \sum_{j=1}^p \sum_{k=1}^q \|V_{jk}\|^2$$
or
$$\min_{V \in \mathbb{R}^{q\times p}} \sum_{i=1}^p \|X_{\cdot,j} - UV_{\cdot,j}^\top\|_{diag(W_{(\cdot,j)})} + \lambda \sum_{i=1}^p \|V_{\cdot j}\|^2$$
This is indeed the sum of p independent problem:
$$\|X_{\cdot,j} - UV_{\cdot j}^\top\|_{diag(W_{(\cdot,j)})} + \lambda\|V_{\cdot j}\|^2$$
which is exactly the equation proposed if we denote $x_j = X_{\cdot,J}$

□