

# Probabilistic Matrix Factorization

Lorenzo Caputi

November 2025

## 1 MAP estimation and trace–norm regularization

### Introduction

The goal of this section is to show that the maximum a posteriori (MAP) estimator in probabilistic matrix factorization (PMF) coincides with the solution of a convex optimization problem involving the trace norm. This equivalence provides a transparent link between the nonconvex landscape of the MAP objective and the convex geometry of trace–norm regularization. It also yields a closed–form characterization of the MAP estimator in terms of the singular value decomposition (SVD) of the data matrix.

We proceed by first expressing the MAP loss in the latent factors  $U, V$ , then using a variational identity to reformulate it as a convex problem in their product  $Z = UV^\top$ . Finally, we solve this convex problem explicitly, obtaining the MAP estimator as a soft–thresholded SVD of  $Y$ .

The connection between MAP estimation and trace–norm regularization has been discussed in both optimization and Bayesian learning literatures. In optimization, trace–norm penalties were introduced in the context of *maximum–margin matrix factorization* (Srebro, Rennie, and Jaakkola, 2005) and further analyzed by Mazumder, Hastie, and Tibshirani (2010) in the context of convex spectral regularization. From a Bayesian perspective, Nakajima et al. (2011) showed that the MAP estimator of a Gaussian latent factor model yields exactly the same shrinkage of singular values as the solution of a trace–norm regularized least squares problem, providing a probabilistic interpretation of the soft–thresholding rule. The argument developed below formalizes this equivalence from first principles, following a deterministic optimization route.

### The MAP loss

Let  $Y \in \mathbb{R}^{n \times p}$  be the observed data matrix, with  $p \geq n$ , and let its rank be  $R$ . We write its singular value decomposition as

$$Y = F \Sigma G^\top, \quad \Sigma = \text{diag}(\sigma_1, \dots, \sigma_R), \quad \sigma_1 \geq \dots \geq \sigma_R > 0.$$

In the PMF model, each row of  $Y$  is represented by a latent vector  $\mathbf{u}_i^\top \in \mathbb{R}^r$  and each column by  $\mathbf{v}_j^\top \in \mathbb{R}^r$ , so that  $Y \approx UV^\top$ . We assume unit observation variance and isotropic Gaussian priors

$$U_{ik} \sim \mathcal{N}(0, \lambda^{-1}), \quad V_{jk} \sim \mathcal{N}(0, \lambda^{-1}),$$

with common precision parameter  $\lambda > 0$ .

The MAP estimator  $(U^*, V^*)$  minimizes the *MAP loss*

$$L(U, V) = \frac{1}{2} \|Y - UV^\top\|_F^2 + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2). \quad (1)$$

The second term acts as a quadratic prior term that enforces smoothness and prevents overfitting. This is a nonconvex problem in  $U$  and  $V$ . Our aim is to show that its global minimum coincides with the solution of a convex optimization problem in the product  $Z = UV^\top$ .

### From factor penalties to the nuclear norm

We start with a classical inequality linking the Frobenius norms of  $U, V$  and the nuclear norm of their product.

**Lemma 1** (Block-PSD inequality). *Let*

$$M = \begin{pmatrix} A & Z \\ Z^\top & B \end{pmatrix} \succeq 0, \quad A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{p \times p}, \quad Z \in \mathbb{R}^{n \times p}.$$

*Then*

$$\text{tr}(A) + \text{tr}(B) \geq 2\|Z\|_*.$$

*Proof.* Since  $M \succeq 0$ , it is symmetric; hence the diagonal blocks are symmetric:  $A = A^\top$  and  $B = B^\top$ . Take the compact SVD  $Z = U\Sigma V^\top$  with  $U = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{n \times r}$ ,  $V = [\mathbf{v}_1, \dots, \mathbf{v}_r] \in \mathbb{R}^{p \times r}$ ,  $U^\top U = V^\top V = I_r$ , and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$  with  $\sigma_i > 0$ .

For any  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{y} \in \mathbb{R}^p$ ,

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}^\top \begin{pmatrix} A & Z \\ Z^\top & B \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{x}^\top A \mathbf{x} + \mathbf{y}^\top B \mathbf{y} + 2\mathbf{x}^\top Z \mathbf{y} \geq 0.$$

Choosing  $\mathbf{x} = \mathbf{u}_i$  and  $\mathbf{y} = -\mathbf{v}_i$  yields

$$\mathbf{u}_i^\top A \mathbf{u}_i + \mathbf{v}_i^\top B \mathbf{v}_i \geq 2 \mathbf{u}_i^\top Z \mathbf{v}_i. \quad (2)$$

**Identify the singular values.** Since  $Z = U\Sigma V^\top$ ,

$$\mathbf{u}_i^\top Z \mathbf{v}_i = \mathbf{u}_i^\top U \Sigma V^\top \mathbf{v}_i = \mathbf{e}_i^\top \Sigma \mathbf{e}_i = \sigma_i.$$

Summing (2) over  $i = 1, \dots, r$  gives

$$\sum_{i=1}^r \mathbf{u}_i^\top A \mathbf{u}_i + \sum_{i=1}^r \mathbf{v}_i^\top B \mathbf{v}_i \geq 2 \sum_{i=1}^r \sigma_i = 2\|Z\|_*. \quad (3)$$

**Trace rewrite and monotonicity.** With  $U = [\mathbf{u}_1, \dots, \mathbf{u}_r]$  and  $V = [\mathbf{v}_1, \dots, \mathbf{v}_r]$ ,

$$\sum_{i=1}^r \mathbf{u}_i^\top A \mathbf{u}_i = \text{tr}(U^\top A U) = \text{tr}(A U U^\top), \quad \sum_{i=1}^r \mathbf{v}_i^\top B \mathbf{v}_i = \text{tr}(V^\top B V) = \text{tr}(B V V^\top).$$

Let  $P := U U^\top$  and  $Q := V V^\top$ . These are orthogonal projectors, hence  $0 \preceq P \preceq I_n$  and  $0 \preceq Q \preceq I_p$ . By trace monotonicity: if  $0 \preceq X \preceq Y$  and  $W \succeq 0$ , then  $\text{tr}(WX) \leq \text{tr}(WY)$ . Indeed, since  $Y - X \succeq 0$ ,

$$\text{tr}(WY) - \text{tr}(WX) = \text{tr}(W^{1/2}(Y - X)W^{1/2}) \geq 0,$$

because  $W^{1/2}(Y - X)W^{1/2} \succeq 0$ . Applying this with  $(X, Y) = (P, I_n)$  for  $A$  and  $(X, Y) = (Q, I_p)$  for  $B$  yields

$$\text{tr}(AP) \leq \text{tr}(A), \quad \text{tr}(BQ) \leq \text{tr}(B).$$

Combining these bounds with (3),

$$\text{tr}(A) + \text{tr}(B) \geq \text{tr}(AP) + \text{tr}(BQ) = \sum_{i=1}^r \mathbf{u}_i^\top A \mathbf{u}_i + \sum_{i=1}^r \mathbf{v}_i^\top B \mathbf{v}_i \geq 2\|Z\|_*,$$

which proves the claim.  $\square$

This yields a convenient variational representation of the nuclear norm.

**Lemma 2** (Variational characterization of the nuclear norm). *For any  $Z \in \mathbb{R}^{n \times p}$ ,*

$$\|Z\|_* = \min_{Z=UV^\top} \frac{1}{2}(\|U\|_F^2 + \|V\|_F^2),$$

and the minimum is achieved by balanced SVD factors.

*Proof.* Fix any factorization  $Z = UV^\top$  with  $U \in \mathbb{R}^{n \times r}$ ,  $V \in \mathbb{R}^{p \times r}$ . Consider the block matrix

$$\begin{pmatrix} UU^\top & Z \\ Z^\top & VV^\top \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} \begin{pmatrix} U^\top & V^\top \end{pmatrix} \succeq 0.$$

Applying Lemma 1 with  $A = UU^\top$ ,  $B = VV^\top$ , and the off-diagonal block  $Z$  yields

$$\text{tr}(UU^\top) + \text{tr}(VV^\top) \geq 2\|Z\|_*.$$

Since  $\text{tr}(UU^\top) = \|U\|_F^2$  and  $\text{tr}(VV^\top) = \|V\|_F^2$ , we obtain

$$\frac{1}{2}(\|U\|_F^2 + \|V\|_F^2) \geq \|Z\|_*.$$

To show that equality can be achieved, take the compact SVD  $Z = \tilde{U}\Sigma\tilde{V}^\top$ , with  $\tilde{U}^\top\tilde{U} = \tilde{V}^\top\tilde{V} = I_r$  and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$ ,  $\sigma_i > 0$ . Set

$$U := \tilde{U}\Sigma^{1/2}, \quad V := \tilde{V}\Sigma^{1/2}.$$

Then  $UV^\top = \tilde{U}\Sigma\tilde{V}^\top = Z$  and

$$\frac{1}{2}(\|U\|_F^2 + \|V\|_F^2) = \frac{1}{2}(\text{tr}(\Sigma) + \text{tr}(\Sigma)) = \text{tr}(\Sigma) = \|Z\|_*.$$

Hence the minimum equals  $\|Z\|_*$ , and it is achieved by the balanced SVD factors.  $\square$

Intuitively, Lemma 2 shows that the sum of the squared norms of  $U$  and  $V$  acts as the tightest possible upper bound for the nuclear norm of  $UV^\top$ . We will now use this identity to reformulate the MAP loss  $L(U, V)$  as a function of  $Z$  only.

### The convex reformulation

Fix  $Z = UV^\top$ . By Lemma 2, for this fixed product the second term in  $L(U, V)$  satisfies

$$\min_{U, V: UV^\top = Z} \frac{1}{2}(\|U\|_F^2 + \|V\|_F^2) = \|Z\|_*.$$

Substituting this into (1) yields a convex function of  $Z$ :

$$M(Z) = \frac{1}{2} \|Y - Z\|_F^2 + \lambda \|Z\|_*. \tag{4}$$

**Proposition 1** (Equivalence of losses).

$$\min_{U, V} L(U, V) = \min_Z M(Z).$$

Moreover, if  $(U^*, V^*)$  minimizes  $L$ , then  $Z^* = U^*V^{*\top}$  minimizes  $M$ ; conversely, if  $Z^*$  minimizes  $M$  and  $Z^* = \tilde{F}\Sigma^*\tilde{G}^\top$  is an SVD, then any balanced factorization

$$U^* = \tilde{F}(\Sigma^*)^{1/2}, \quad V^* = \tilde{G}(\Sigma^*)^{1/2}$$

minimizes  $L$ .

*Proof.* We first establish that  $\min_{U, V} L(U, V) \geq \min_Z M(Z)$ .

For any factorization  $U, V$  with  $Z = UV^\top$ , by Lemma 2 we have

$$\frac{1}{2}(\|U\|_F^2 + \|V\|_F^2) \geq \|Z\|_*,$$

with equality achieved by balanced SVD factors. Therefore,

$$L(U, V) = \frac{1}{2}\|Y - UV^\top\|_F^2 + \frac{\lambda}{2}(\|U\|_F^2 + \|V\|_F^2) \geq \frac{1}{2}\|Y - Z\|_F^2 + \lambda\|Z\|_* = M(Z).$$

Taking the minimum over all  $U, V$  on the left and noting that every  $Z$  can be achieved by some factorization  $UV^\top$ , we get

$$\min_{U, V} L(U, V) \geq \min_Z M(Z).$$

For the reverse inequality, let  $Z^*$  minimize  $M(Z)$ . Take the SVD  $Z^* = \tilde{F}\Sigma^*\tilde{G}^\top$  and define balanced factors

$$U^* = \tilde{F}(\Sigma^*)^{1/2}, \quad V^* = \tilde{G}(\Sigma^*)^{1/2}.$$

Then  $U^*V^{*\top} = Z^*$  and by Lemma 2, equality holds in the variational characterization:

$$\frac{1}{2}(\|U^*\|_F^2 + \|V^*\|_F^2) = \|Z^*\|_*.$$

Therefore,

$$L(U^*, V^*) = \frac{1}{2}\|Y - Z^*\|_F^2 + \lambda\|Z^*\|_* = M(Z^*) = \min_Z M(Z).$$

This shows  $\min_{U, V} L(U, V) \leq \min_Z M(Z)$ .

Combining both inequalities gives  $\min_{U, V} L(U, V) = \min_Z M(Z)$ .

For the second part, if  $(U^*, V^*)$  minimizes  $L$ , then by the first part and the fact that  $L(U^*, V^*) = \min_{U, V} L(U, V) = \min_Z M(Z)$ , we must have  $M(Z^*) = \min_Z M(Z)$  where  $Z^* = U^*V^{*\top}$ . Hence  $Z^*$  minimizes  $M$ .

Conversely, if  $Z^*$  minimizes  $M$  and we take balanced factors as constructed above, then  $L(U^*, V^*) = M(Z^*) = \min_Z M(Z) = \min_{U, V} L(U, V)$ , so  $(U^*, V^*)$  minimizes  $L$ .  $\square$

This establishes the desired bridge: the nonconvex MAP problem  $L(U, V)$  and the convex trace-norm problem  $M(Z)$  have the same global minimum value and the same optimal product  $Z^*$ .

## Solving the convex problem

We now derive the explicit form of the minimizer of  $M(Z)$  in (4). To do so, we need a standard spectral inequality that allows us to align the singular vectors of  $Y$  and  $Z$ .

**Lemma 3** (von Neumann's trace inequality). *For any matrices  $A, B$  with singular values  $\{\alpha_i\}$  and  $\{\beta_i\}$  in non-increasing order,*

$$\langle A, B \rangle \leq \sum_i \alpha_i \beta_i,$$

where  $\langle A, B \rangle = \text{tr}(A^\top B)$  and  $\|A\|_F^2 = \langle A, A \rangle$ . Equality holds if and only if the singular vectors of  $A$  and  $B$  are aligned.

*Proof.* Let  $A, B \in \mathbb{R}^{n \times p}$  and let  $q = \min\{n, p\}$ . Express  $A$  and  $B$  using their singular value decompositions:  $A = U_A \Sigma_A V_A^\top$  and  $B = U_B \Sigma_B V_B^\top$ , where  $U_A, V_A, U_B, V_B$  are orthogonal matrices and  $\Sigma_A = \text{diag}(\alpha_1, \dots, \alpha_q)$  and  $\Sigma_B = \text{diag}(\beta_1, \dots, \beta_q)$  contain the ordered singular values  $\alpha_1 \geq \dots \geq \alpha_q$  and  $\beta_1 \geq \dots \geq \beta_q$ , respectively.

The trace of their product can be written as

$$\text{tr}(A^\top B) = \text{tr}(V_A \Sigma_A U_A^\top U_B \Sigma_B V_B^\top).$$

By the cyclic property of the trace,

$$\text{tr}(A^\top B) = \text{tr}(\Sigma_A M \Sigma_B N),$$

where  $M = U_A^\top U_B$  and  $N = V_B^\top V_A$  are orthogonal matrices.

Expanding the trace:

$$\text{tr}(\Sigma_A M \Sigma_B N) = \sum_{i=1}^q \sum_{j=1}^q \alpha_i \beta_j (M \Sigma_B N)_{ij}.$$

Since  $M$  and  $N$  are orthogonal, the matrix with entries  $|(M \Sigma_B N)_{ij}|$  is dominated by a doubly stochastic matrix (entries are nonnegative and each row and column sums to at most 1). By the Birkhoff–von Neumann theorem, any doubly stochastic matrix can be expressed as a convex combination of permutation matrices. Therefore,

$$|\text{tr}(A^\top B)| = \left| \sum_{i=1}^q \sum_{j=1}^q \alpha_i \beta_j (M \Sigma_B N)_{ij} \right| \leq \max_{\pi} \sum_{i=1}^q \alpha_i \beta_{\pi(i)},$$

where the maximum is over all permutations  $\pi$  of  $\{1, \dots, q\}$ . The maximum is attained when the largest  $\alpha_i$  values are paired with the largest  $\beta_j$  values, giving

$$\max_{\pi} \sum_{i=1}^q \alpha_i \beta_{\pi(i)} = \sum_{i=1}^q \alpha_i \beta_i.$$

Therefore,

$$\text{tr}(A^\top B) \leq \sum_{i=1}^q \alpha_i \beta_i,$$

with equality if and only if the singular vectors of  $A$  and  $B$  are aligned such that the singular values are matched in decreasing order.  $\square$

Applying Lemma 3 to the quadratic term  $\|Y - Z\|_F^2$ , one sees that the optimal  $Z$  must share

its singular vectors with  $Y$ . The problem therefore reduces to a separable minimization over the singular values  $s_i \geq 0$  of  $Z$ :

$$\min_{s_i \geq 0} \sum_i \left[ \frac{1}{2}(\sigma_i - s_i)^2 + \lambda s_i \right].$$

Each term is minimized independently at

$$s_i^* = (\sigma_i - \lambda)_+ := \max(\sigma_i - \lambda, 0).$$

This operation is known in the optimization literature as singular value shrinkage (Mazumder et al., 2010).

**Proposition 2** (Closed-form solution). *Let  $Y = F \Sigma G^\top$ . The unique minimizer of  $M(Z)$  is*

$$Z^* = F \operatorname{diag}((\sigma_i - \lambda)_+) G^\top.$$

This solution has a simple interpretation: the MAP estimate  $Z^*$  retains the singular vectors of  $Y$  and shrinks each singular value by  $\lambda$ , truncating to zero when  $\sigma_i \leq \lambda$ .

## Recovering the MAP factors

Since the minimizers of  $L$  and  $M$  coincide through  $Z^* = U^* V^{*\top}$ , any balanced SVD factorization of  $Z^*$  gives a valid MAP pair:

$$U^* = F \operatorname{diag}((\sigma_i - \lambda)_+^{1/2}), \quad V^* = G \operatorname{diag}((\sigma_i - \lambda)_+^{1/2}).$$

The pair  $(U^*, V^*)$  is not unique, since for any orthogonal matrix  $Q$ ,  $(U^* Q, V^* Q)$  yields the same product  $Z^*$  and the same loss value.

## Discussion and relation to previous work

We have shown that the MAP estimator in probabilistic matrix factorization with isotropic Gaussian priors is equivalent to the solution of a convex trace–norm regularization problem. The equivalence can be summarized succinctly as

$$\boxed{\min_{U,V} \left[ \frac{1}{2} \|Y - UV^\top\|_F^2 + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2) \right] = \min_Z \left[ \frac{1}{2} \|Y - Z\|_F^2 + \lambda \|Z\|_* \right].}$$

The minimizer  $Z^* = F \operatorname{diag}((\sigma_i - \lambda)_+) G^\top$  gives the posterior mode of the reconstructed matrix. The prior precision  $\lambda$  acts as a singular-value shrinkage parameter, promoting low rank whenever  $\lambda$  exceeds some of the  $\sigma_i$ .

This equivalence makes explicit that regularization in matrix factorization operates by shrinking the singular values of the data matrix. Small singular values typically capture noise or idiosyncratic variation, while large ones represent systematic latent structure. Hence, the MAP penalty  $\lambda$  can

be interpreted as a threshold that filters out uninformative directions in the data. In this sense, the MAP estimator trades data fidelity for simplicity in exactly the same way as the convex spectral regularizer of Srebro and Jaakkola (2005).

It is interesting to note that Theorem 1 in Nakajima et al. (2011) derives the same form of estimator using a fully Bayesian argument. In that setting, the shrinkage of singular values arises as a consequence of Gaussian priors on the factor columns, leading again to a soft-thresholded SVD of  $Y$ .

This equivalence forms the analytical foundation for our study of the loss landscape in the next section, where we examine the nature of stationary points of  $L(U, V)$  and the role of the soft-thresholded singular values in determining the global minima.

## 2 Rank-1 ALS dynamics for the MAP objective

**Proposition 3** (Invariance and scalar recursion). *Fix  $i \in \{1, \dots, n\}$  and assume  $\mathbf{v}_0 = \gamma_0 \mathbf{g}_i$  with  $\gamma_0 > 0$ . Then the ALS iterates remain in the singular pair  $(\mathbf{f}_i, \mathbf{g}_i)$  and admit the forms*

$$\mathbf{u}_k = \gamma_{2k-1} \mathbf{f}_i, \quad \mathbf{v}_k = \gamma_{2k} \mathbf{g}_i \quad (k \geq 0),$$

where the nonnegative sequence  $(\gamma_t)_{t \geq 0}$  satisfies the one-dimensional recurrence

$$\boxed{\gamma_{t+1} = g(\gamma_t) := \frac{\sigma_i \gamma_t}{\lambda + \gamma_t^2}, \quad t = 0, 1, 2, \dots}$$

*Proof.* From the SVD identities  $Y \mathbf{g}_i = \sigma_i \mathbf{f}_i$  and  $Y^\top \mathbf{f}_i = \sigma_i \mathbf{g}_i$  we obtain: if  $\mathbf{v} = \gamma \mathbf{g}_i$ , then

$$\mathbf{u} \leftarrow Y \mathbf{v} (\mathbf{v}^\top \mathbf{v} + \lambda)^{-1} = \frac{\sigma_i \gamma}{\lambda + \gamma^2} \mathbf{f}_i;$$

and if  $\mathbf{u} = \gamma \mathbf{f}_i$ , then

$$\mathbf{v} \leftarrow Y^\top \mathbf{u} (\mathbf{u}^\top \mathbf{u} + \lambda)^{-1} = \frac{\sigma_i \gamma}{\lambda + \gamma^2} \mathbf{g}_i.$$

Thus each half-step rescales by the same scalar map  $g(\cdot)$  and preserves the directions  $(\mathbf{f}_i, \mathbf{g}_i)$ . Writing the successive half-steps as  $\gamma_{t+1} = g(\gamma_t)$  gives the claim.  $\square$

**Proposition 4** (Fixed points, monotone convergence, and local rate). *Consider the recursion  $\gamma_{t+1} = g(\gamma_t) = \sigma_i \gamma_t / (\lambda + \gamma_t^2)$  with  $\gamma_0 > 0$ .*

(a) **Fixed points.** *The fixed points are  $\gamma = 0$  and, if  $\sigma_i > \lambda$ ,  $\gamma^* = \sqrt{\sigma_i - \lambda}$ .*

(b) **Global monotone convergence.**

- If  $\sigma_i \leq \lambda$ , then  $\gamma_t \downarrow 0$ .
- If  $\sigma_i > \lambda$ , then  $\gamma_t \rightarrow \gamma^* = \sqrt{\sigma_i - \lambda}$  monotonically: it increases when  $\gamma_t < \gamma^*$  and decreases when  $\gamma_t > \gamma^*$ .

(c) **Local linear rate.** If  $\sigma_i > \lambda$ , convergence to  $\gamma^*$  is linear with factor  $|g'(\gamma^*)| = |1 - 2\lambda/\sigma_i| < 1$ . Equivalently, for some  $C > 0$  and all large  $t$ ,

$$|\gamma_t - \gamma^*| \leq C |1 - 2\lambda/\sigma_i|^t.$$

*Proof.* (a) Solving  $g(\gamma) = \gamma$  gives  $\gamma = \sigma_i\gamma/(\lambda + \gamma^2)$ , i.e.  $\gamma(\lambda + \gamma^2 - \sigma_i) = 0$ . Hence  $\gamma = 0$  or  $\gamma^2 = \sigma_i - \lambda$ , the latter only if  $\sigma_i > \lambda$ .

(b) Let  $h(\gamma) = g(\gamma) - \gamma = \gamma(\sigma_i - \lambda - \gamma^2)/(\lambda + \gamma^2)$ . For  $\gamma > 0$ ,  $\text{sign } h(\gamma) = \text{sign}(\sigma_i - \lambda - \gamma^2)$ . If  $\sigma_i \leq \lambda$ , then  $h(\gamma) < 0$  for all  $\gamma > 0$ , so  $\gamma_{t+1} < \gamma_t$  and  $\gamma_t \downarrow 0$ . If  $\sigma_i > \lambda$ , set  $\gamma^* = \sqrt{\sigma_i - \lambda}$ . Then  $h(\gamma) > 0$  for  $\gamma \in (0, \gamma^*)$  and  $h(\gamma) < 0$  for  $\gamma > \gamma^*$ , so  $\gamma_t$  moves monotonically toward  $\gamma^*$  and converges.

(c) The derivative  $g'(\gamma) = \sigma_i(\lambda - \gamma^2)/(\lambda + \gamma^2)^2$  evaluated at  $\gamma^*$  yields  $g'(\gamma^*) = 1 - 2\lambda/\sigma_i$ , whose magnitude is  $< 1$ . Linearizing  $g$  near  $\gamma^*$  gives  $\gamma_{t+1} - \gamma^* \approx g'(\gamma^*)(\gamma_t - \gamma^*)$ , so  $|\gamma_t - \gamma^*| \leq C |1 - 2\lambda/\sigma_i|^t$  for some  $C > 0$ .  $\square$

**Proposition 5** (MAP optimum on the ray and classification). *Restricting the loss  $L(\mathbf{u}, \mathbf{v})$  to the ray  $\mathbf{u} = a\mathbf{f}_i$ ,  $\mathbf{v} = b\mathbf{g}_i$  yields*

$$\min_{a,b \in \mathbb{R}} \frac{1}{2}(\sigma_i - ab)^2 + \frac{\lambda}{2}(a^2 + b^2),$$

whose unique minimizer is  $a = b = \sqrt{(\sigma_i - \lambda)_+}$ . Hence the ALS limit described in Proposition 4 coincides with the MAP optimum in the  $(\mathbf{f}_i, \mathbf{g}_i)$  direction. By the general characterization of the MAP landscape, this stationary point is a global minimum if  $i = 1$ , and a strict saddle if  $i > 1$ .

*Proof.* Along  $\mathbf{u} = a\mathbf{f}_i$ ,  $\mathbf{v} = b\mathbf{g}_i$ , orthonormality of  $\{\mathbf{f}_j, \mathbf{g}_j\}$  gives  $\|Y - \mathbf{u}\mathbf{v}^\top\|_F^2 = \sum_{j \neq i} \sigma_j^2 + (\sigma_i - ab)^2$ , so up to an additive constant the restricted problem is  $\phi(a, b) = \frac{1}{2}(\sigma_i - ab)^2 + \frac{\lambda}{2}(a^2 + b^2)$ . The first-order conditions are  $-(\sigma_i - ab)b + \lambda a = 0$  and  $-(\sigma_i - ab)a + \lambda b = 0$ . Subtracting gives  $(a - b)(\lambda + \sigma_i - ab) = 0$ . The symmetric positive solution has  $a = b \geq 0$ , and substituting yields  $a(\lambda + a^2 - \sigma_i) = 0$ , so  $a = b = \sqrt{(\sigma_i - \lambda)_+}$ . Strict convexity of  $\phi$  ensures uniqueness. By the general MAP landscape result for rank- $K$  factorisation, this point is a global minimum if it corresponds to the top singular direction ( $i = 1$ ), and a strict saddle otherwise.  $\square$

**Proposition 6** (Subspace invariance and reduction to one-dimensional dynamics). *Fix  $n \leq p$  and let  $Y \in \mathbb{R}^{n \times p}$  have singular value decomposition  $Y = F\Sigma G^\top$ . Let  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_R) \succ 0$  be a diagonal precision matrix, and consider the MAP objective*

$$L(U, V) = \frac{1}{2}\|Y - UV^\top\|_F^2 + \frac{1}{2}(\|U\Lambda^{1/2}\|_F^2 + \|V\Lambda^{1/2}\|_F^2),$$

with alternating ridge updates

$$U \leftarrow YV(V^\top V + \Lambda)^{-1}, \quad V \leftarrow Y^\top U(U^\top U + \Lambda)^{-1}.$$

Select any index set  $\mathcal{I} = \{i_1, \dots, i_R\} \subset \{1, \dots, n\}$  and define the corresponding singular subspace

$$\tilde{F}_R := [\mathbf{f}_{i_1}, \dots, \mathbf{f}_{i_R}], \quad \tilde{G}_R := [\mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_R}], \quad \tilde{\Sigma}_R := \text{diag}(\sigma_{i_1}, \dots, \sigma_{i_R}).$$

Initialize ALS with scaled singular vectors

$$U_0 = \tilde{F}_R \Gamma_0, \quad V_0 = \tilde{G}_R \Gamma_0,$$

where  $\Gamma_0$  is diagonal with positive entries. Define a sequence of diagonal matrices  $\Gamma_k$  such that

$$U_k = \tilde{F}_R \Gamma_{2k-1}, \quad V_k = \tilde{G}_R \Gamma_{2k}, \quad k \geq 0.$$

**Statement.** If the initialization consists of scaled singular vectors as above, then:

- (i) **Invariance of singular directions.** The ALS iterations remain within the chosen singular subspace:

$$U_k \in \text{span}(\tilde{F}_R), \quad V_k \in \text{span}(\tilde{G}_R),$$

and the coefficient matrices  $\Gamma_k$  remain diagonal for all  $k$ . In other words, ALS never changes direction: it only rescales each selected singular vector.

- (ii) **Coordinate-wise dynamics.** The sequence of diagonal matrices  $(\Gamma_k)$  evolves according to

$$\Gamma_{t+1} = \tilde{\Sigma}_R \Gamma_t (\Gamma_t^2 + \Lambda)^{-1}, \quad t = 0, 1, 2, \dots,$$

where all matrices are diagonal and products are taken elementwise on the diagonals. Writing  $\Gamma_t = \text{diag}(\gamma_{t,1}, \dots, \gamma_{t,R})$ , each coordinate satisfies

$$\gamma_{t+1,r} = \frac{\sigma_{i_r} \gamma_{t,r}}{\lambda_r + \gamma_{t,r}^2}, \quad r = 1, \dots, R.$$

Hence each scale  $\gamma_{t,r}$  evolves independently following the same scalar map studied in Proposition 3.

- (iii) **Interpretation.** Along every selected singular direction  $(\mathbf{f}_{i_r}, \mathbf{g}_{i_r})$ , ALS behaves exactly as in the rank-one analysis: it moves only along that ray, converging to the stationary scale

$$\gamma_r^* = \begin{cases} 0, & \text{if } \sigma_{i_r} \leq \lambda_r, \\ \sqrt{\sigma_{i_r} - \lambda_r}, & \text{if } \sigma_{i_r} > \lambda_r. \end{cases}$$

Along this ray the MAP loss is locally convex, and ALS converges as if approaching a local minimum. The saddle nature of non-top subspaces appears only when moving transversally—i.e., towards a direction involving a larger singular value—exactly as shown in the saddlepoint analysis.

*Proof.* Write  $U = \tilde{F}_R A$  and  $V = \tilde{G}_R B$  with  $A, B \in \mathbb{R}^{R \times R}$ . Using orthogonality of the singular vectors,

$$YV = \tilde{F}_R \tilde{\Sigma}_R B, \quad Y^\top U = \tilde{G}_R \tilde{\Sigma}_R A.$$

Substituting into the ALS updates gives

$$A \leftarrow \tilde{\Sigma}_R B (B^\top B + \Lambda)^{-1}, \quad B \leftarrow \tilde{\Sigma}_R A (A^\top A + \Lambda)^{-1}.$$

If  $A$  and  $B$  are diagonal, then  $B^\top B + \Lambda$  and  $A^\top A + \Lambda$  are diagonal and invertible, and the right-hand sides remain diagonal. Thus diagonality (and hence the alignment with  $\tilde{F}_R, \tilde{G}_R$ ) is preserved. With the parity convention  $\Gamma_{2k-1} = A_k$  and  $\Gamma_{2k} = B_k$ , the updates reduce to

$$\Gamma_{t+1} = \tilde{\Sigma}_R \Gamma_t (\Gamma_t^2 + \Lambda)^{-1},$$

which maps diagonal matrices to diagonal matrices and proves (i)–(ii). Taking the diagonal entries gives  $\gamma_{t+1,r} = \sigma_{i_r} \gamma_{t,r} / (\lambda_r + \gamma_{t,r}^2)$ , which is exactly the scalar recurrence analyzed in Proposition 3. Hence each coordinate inherits the same dynamics, fixed points, and convergence behavior described in Propositions 3 and 4. Since along each selected ray the objective reduces to the one-dimensional function  $\frac{1}{2}(\sigma_{i_r} - \gamma^2)^2 + \frac{\lambda_r}{2}(2\gamma^2)$ , the loss is locally convex in that direction, while saddles arise only when moving towards other singular directions (where a larger singular value decreases the loss).  $\square$

## A Stationary Points and the Grassmann Manifold in the ML/PCA Case

### A.1 Interpretation of PCA as optimization over the Grassmann Manifold

In the maximum likelihood (or equivalently, PCA) case, the loss function is

$$L(U, V) = \frac{1}{2} \|Y - UV^\top\|_F^2,$$

and its stationary points satisfy

$$YV = U(V^\top V), \quad Y^\top U = V(U^\top U).$$

It is known that every stationary point  $(U^*, V^*)$  can be written as

$$V^* = G_R C,$$

where  $G_R$  contains a subset of  $R$  right singular vectors of  $Y$ , and  $C \in \mathbb{R}^{R \times R}$  is any invertible matrix.

**Invariance and equivalence classes.** The value of the loss at these stationary points depends only on the choice of  $G_R$ , not on  $C$ . In fact, if we replace  $(U, V)$  by

$$(U', V') = (UC^{-\top}, VC),$$

for any invertible  $C$ , the product  $UV^\top$ —and hence the loss  $L(U, V)$ —remains unchanged:

$$(UC^{-\top})(VC)^\top = UC^{-\top}C^\top V^\top = UV^\top.$$

Therefore, stationary points are not isolated, but form *equivalence classes* under this transformation. Each class consists of all matrices that can be obtained by multiplying a given  $V$  by an invertible  $C$ .

**Equivalence classes as subspaces.** Take any two matrices  $V_0$  and  $V_1 = V_0C$  for some invertible  $C$ . The columns of  $V_1$  are linear combinations of the columns of  $V_0$ :

$$\text{span}(V_1) = \text{span}(V_0).$$

Thus, all matrices in the same equivalence class share the same column space, and differ only by a change of basis within that subspace. In this sense, the object identified by a stationary point is not a specific matrix  $V^*$ , but rather the subspace spanned by its columns.

**The Grassmann manifold perspective.** Each stationary point therefore corresponds to an  $R$ -dimensional subspace of  $\mathbb{R}^n$ , i.e. an element of the Grassmann manifold

$$\text{Gr}(R, n) = \{ \mathcal{V} \subset \mathbb{R}^n \mid \dim(\mathcal{V}) = R \}.$$

From this viewpoint, all stationary points of the ML/PCA objective lie on  $\text{Gr}(R, n)$ . The optimization problem is thus naturally defined on the Grassmann manifold: what is being estimated is a subspace—the principal subspace—rather than a particular basis for it.

## A.2 A saddlepoint argument based on the Grassmann Manifold

Let  $Y \in \mathbb{R}^{n \times p}$  and  $S := Y^\top Y \in \mathbb{R}^{p \times p}$ . In the ML/PCA case

$$L(U, V) = \frac{1}{2} \|Y - UV^\top\|_F^2.$$

For any fixed full-column  $V \in \mathbb{R}^{p \times R}$ , the  $U$ -minimizer is  $U^*(Y, V) = YV(V^\top V)^{-1}$ , hence the reduced objective is

$$\min_U L(U, V) = \frac{1}{2} \|Y - YV(V^\top V)^{-1}V^\top\|_F^2 = \frac{1}{2} \|Y(I - P_V)\|_F^2,$$

where  $P_V := V(V^\top V)^{-1}V^\top$  is the orthogonal projector onto  $\text{span}(V)$ . Using  $\|Y(I - P_V)\|_F^2 = \text{tr}(S) - \text{tr}(P_V S)$ , we get

$$\min_U L(U, V) = \frac{1}{2}\text{tr}(S) - \frac{1}{2}\text{tr}(P_V S).$$

Therefore, at fixed rank  $R$ , minimizing  $L$  is equivalent to maximizing

$$F(\mathcal{V}) := \text{tr}(P_{\mathcal{V}} S) \quad \text{over } \mathcal{V} \in \text{Gr}(R, p),$$

where  $P_{\mathcal{V}}$  is the orthogonal projector onto the subspace  $\mathcal{V} \subset \mathbb{R}^p$  of dimension  $R$ .

**Stationary subspaces.** Let  $S = G \text{diag}(\mu_1, \dots, \mu_p) G^\top$  with  $\mu_1 \geq \dots \geq \mu_p$ , and  $G = [\mathbf{g}_1 \ \dots \ \mathbf{g}_p]$  orthogonal. By the *spectral theorem* for symmetric matrices, the eigenvectors  $\mathbf{g}_1, \dots, \mathbf{g}_p$  can be chosen to be orthonormal, i.e.  $G^\top G = I_p$ . It is standard that  $\mathcal{V}$  is a stationary point of  $F$  iff  $\mathcal{V}$  is  $S$ -invariant, i.e.  $\mathcal{V} = \text{span}\{\mathbf{g}_i : i \in I\}$  for some index set  $I \subset \{1, \dots, p\}$  with  $|I| = R$ . Equivalently, any stationary  $V$  can be written as  $V = G_I C$  with  $C \in \text{GL}(R)$ . Moreover,

$$F(\mathcal{V}) = \sum_{i \in I} \mu_i,$$

so the (global) maximizers correspond to  $I = \{1, \dots, R\}$  (Eckart–Young).

**Saddle property for non–top- $R$  stationary subspaces.** Fix a stationary  $\mathcal{V} = \text{span}\{\mathbf{g}_i : i \in I\}$  with  $I \neq \{1, \dots, R\}$ . Then there exist indices  $i \in I$  and  $j \notin I$  such that  $\mu_j > \mu_i$ . Consider the 2-plane  $\Pi = \text{span}\{\mathbf{g}_i, \mathbf{g}_j\}$  and define a smooth curve of  $R$ -dimensional subspaces by *rotating*  $\mathbf{g}_i$  toward  $\mathbf{g}_j$  inside  $\Pi$ :

$$\tilde{\mathbf{g}}_i(\theta) = \cos \theta \mathbf{g}_i + \sin \theta \mathbf{g}_j, \quad \mathcal{V}(\theta) = \text{span}(\{\tilde{\mathbf{g}}_i(\theta)\} \cup \{\mathbf{g}_k : k \in I \setminus \{i\}\}).$$

Let  $P(\theta)$  be the projector onto  $\mathcal{V}(\theta)$ . Since  $\{\mathbf{g}_k : k \in I \setminus \{i\}\}$  and  $\tilde{\mathbf{g}}_i(\theta)$  form an orthonormal set, we can express  $P(\theta)$  as a sum of *rank-one projectors*:

$$P(\theta) = \sum_{k \in I \setminus \{i\}} \mathbf{g}_k \mathbf{g}_k^\top + \tilde{\mathbf{g}}_i(\theta) \tilde{\mathbf{g}}_i(\theta)^\top.$$

This follows from the fact that for any orthonormal basis  $\{\mathbf{v}_r\}$  of a subspace, the orthogonal projection onto their span is  $\sum_r \mathbf{v}_r \mathbf{v}_r^\top$ .

Hence

$$F(\mathcal{V}(\theta)) = \text{tr}(P(\theta) S) = \sum_{k \in I \setminus \{i\}} \mathbf{g}_k^\top S \mathbf{g}_k + \tilde{\mathbf{g}}_i(\theta)^\top S \tilde{\mathbf{g}}_i(\theta) = \sum_{k \in I \setminus \{i\}} \mu_k + \mu_i \cos^2 \theta + \mu_j \sin^2 \theta.$$

Taking derivatives gives

$$\frac{d}{d\theta} F(\mathcal{V}(\theta)) \Big|_{\theta=0} = 0, \quad \frac{d^2}{d\theta^2} F(\mathcal{V}(\theta)) \Big|_{\theta=0} = 2(\mu_j - \mu_i) > 0.$$

Therefore, along this admissible (Grassmann) direction the objective  $F$  has *upward* curvature: we can strictly increase  $F$  (and thus strictly decrease the original loss) by moving away from  $\mathcal{V}$ .

To see that  $\mathcal{V}$  is a *saddle* rather than a local maximum, we exhibit a direction of *downward* curvature. If there exists  $\ell \notin I$  with  $\mu_\ell < \mu_i$ , repeat the same construction rotating  $\mathbf{g}_i$  toward  $\mathbf{g}_\ell$ ; the same computation gives  $\frac{d^2}{d\theta^2} F(\mathcal{V}(\theta))|_{\theta=0} = 2(\mu_\ell - \mu_i) < 0$ . If no such  $\ell$  exists, then  $i$  must index the *smallest* eigenvalue among those in  $I$ , and there is some  $r \in I$  with  $\mu_r > \mu_i$ ; rotating  $\mathbf{g}_r$  toward a  $\mathbf{g}_\ell$  with  $\mu_\ell < \mu_r$  again yields a negative curvature direction. In all cases, the Riemannian Hessian of  $F$  at  $\mathcal{V}$  has both positive and negative directions in the tangent space  $T_{\mathcal{V}}\text{Gr}(R, p)$ .

**Conclusion.** Every stationary subspace that does not span the top- $R$  eigenvectors of  $S$  admits ascent and descent directions on  $\text{Gr}(R, p)$ ; hence it is a *saddle point* of  $F$  and, equivalently, of the reduced PCA objective  $\min_U L(U, V)$ . Only the top- $R$  subspace is a (global) maximizer of  $F$  and thus a (global) minimizer of  $L$ .