STAT 538: Statistical Learning: Modeling, Prediction, And Computing **Winter 2023**

Lecture 16: Rank-Sparsity Matrix Decomposition

Lecturer: Armeen Taeb *Scribe: Yikun Zhang*

Parts of the notes are based on Chandrasekaran et al. [2009, 2011].

Setting: Let $C = A^* + B^*$ with $A^* \in \mathbb{R}^{n\times n}$ being a sparse matrix and $B^* \in \mathbb{R}^{n\times n}$ a low-rank matrix, where both A^* and B^* are unknown. In this notes, we restrict ourselves to square matrices in $\mathbb{R}^{n\times n}$, but the analysis can be extended to rectangular matrices $\mathbb{R}^{n_1\times n_2}$ if we simply replace n by $\max\{n_1, n_2\}$.

Goal: Given C, we want to recover A^* and B^* without any prior information about the sparsity pattern of A^* or the rank/singular vectors of B^*.

Solution: Consider the following optimization problem:

$$\begin{aligned} &\arg\min_{A,B} \left[\gamma \left|\left|A\right|\right|_1 + \left|\left|B\right|\right|_\star\right] \\ &\quad \text{subject to } A + B = C. \end{aligned} \tag{1}$$

Here, $||A||_1 = \sum_{i,j} |A_{ij}|$ is the elementwise L_1-norm of a matrix A, $||B||_\star = \sum_k \sigma_k(B)$ is the nuclear norm, which is the sum of the singular values of B, and γ is a tuning parameter that provides a trade-off between the low-rank and sparse components.

Remark 1. *This optimization problem (1) is convex and can be written as a semi-definite program (SDP; Vandenberghe and Boyd 1996), for which there exist polynomial-time general- purpose solvers; see Appendix A in Chandrasekaran et al. [2011]. Under a mild tightening of the conditions for fundamental identifiability, the minimizer of (1) is unique and recover A^*, B^*. Essentially, these conditions require that the sparse matrix does not have support concentrated within a single row/column, while the low-rank matrix does not have row/column spaces closely aligned with the coordinate axes [Chandrasekaran et al., 2009].*

Notations: We begin by introducing several algebraic varieties[1]. The set of rank-constrained matrices is defined as:

$$\mathcal{P}(k) = \left\{M \in \mathbb{R}^{n\times n} : \mathrm{rank}(M) \le k\right\}.$$

This is an algebraic variety with dimension $k(2n-k) = n^2 - (n-k)^2$, since it can be defined through the vanishing of all $(k+1)\times(k+1)$ minors of the matrix M. Let $M = UDV^T \in \mathbb{R}^{n\times n}$ be the singular value decomposition of M with $U, V \in \mathbb{R}^{n\times k}$ and $\mathrm{rank}(M) = k$. The tangent space at M is defined as:

$$T(M) = \left\{UX^T + YV^T : X, Y \in \mathbb{R}^{n\times n}\right\},$$

which consists of the span of all matrices with either the same row space as M or the same column space as M. We also define

$$\Omega(M) = \left\{N \in \mathbb{R}^{n\times n} : \mathrm{support}(N) \subseteq \mathrm{support}(M)\right\},$$

which is the tangent space of $\{M \in \mathbb{R}^{n\times n} : |\mathrm{support}(M)| \le m\}$. Consider the following two quantities:

$$\xi(M) = \max_{N\in T(M), ||N||_2 \le 1} ||N||_\infty$$

[1] Recall that an algebraic variety is defined as the zero set of a system of polynomial equations [Hartshorne, 2013].

which will be small when (appropriately scaled) elements of the tangent space $T(M)$ are "diffuse" (*i.e.*, these elements are not too sparse), and

$$\mu(M) = \max_{N \in \Omega(M), ||N||_\infty \leq 1} ||N||_2$$

which will be small when the spectrum of any matrix in $\Omega(M)$ is "diffuse" (*i.e.*, the singular values of these elements are not too large). Here, $||\cdot||_\infty$ denotes the largest entry in magnitude and $||\cdot||_2$ is the spectral norm (*i.e.*, the largest singular value).

Remark 2. *One can show that*

$$\deg_{\min}(M) \leq \mu(M) \leq \deg_{\max}(M),$$

where $\deg_{\max}(M)$ *is the maximum number of nonzero entries per row/column and* $\deg_{\min}(M)$ *is the minimum number of nonzero entries per row/column; see Proposition 3 in Chandrasekaran et al. [2011]. Analogously, we can bound* $\xi(M)$ *as:*

$$\mathrm{inc}(M) \leq \xi(M) \leq 2 \cdot \mathrm{inc}(M),$$

where $\mathrm{inc}(M) = \max\{\beta(\textit{row-space}(M)), \beta(\textit{column-space}(M))\}$ *is the incoherence of the row/column spaces of a matrix* $M \in \mathbb{R}^{n\times n}$ *with* $\beta(S) = \max_i ||P_S e_i||_2$ *as the incoherence of a subspace* $S \subset \mathbb{R}^n$*. Here,* $\{e_1, ..., e_n\}$ *is the standard basis of* $\mathbb{R}^n$*,* P_S *denotes the projection onto the subspace* S*, and* $||\cdot||_2$ *is the vector* ℓ_2*-norm.*

1 Basic Properties

Proposition 1. *If* $\mu(A^*)\xi(B^*) < 1$ *for two matrices* $A^*, B^* \in \mathbb{R}^{n\times n}$*, then* $\Omega(A^*) \cap T(B^*) = \{0\}$*.*

We may choose γ properly to have $\mu(A^*)\xi(B^*) < 1/6$, which guarantees the recoveries of A^* and B^*. To establish Proposition 1, we leverage the following lemma.

Lemma 2. $\max_{N \in T(B^*), ||N||_2 \leq 1} \left|\left|P_{\Omega(A^*)}(N)\right|\right|_2 \leq \mu(A^*) \cdot \xi(B^*)$*, where* $P_{\Omega(A^*)}(N)$ *is the projection of* N *on the space* $\Omega(A^*)$*.*

Proof of Lemma 2. We have the following sequence of inequalities:

$$\begin{aligned}
\max_{N \in T(B^*), ||N||_2 \leq 1} \left|\left|P_{\Omega(A^*)}(N)\right|\right|_2 &\leq \max_{N \in T(B^*), ||N||_2 \leq 1} \mu(A^*) \left|\left|P_{\Omega(A^*)}(N)\right|\right|_\infty \\
&\leq \max_{N \in T(B^*), ||N||_2 \leq 1} \mu(A^*) ||N||_\infty \\
&= \mu(A^*) \cdot \xi(B^*),
\end{aligned}$$

where the first inequality follows from the definition of $\mu(A^*)$ as $P_{\Omega(A^*)}(N) \in \Omega(A^*)$ and the second inequality is due to $\left|\left|P_{\Omega(A^*)}(N)\right|\right|_\infty \leq ||N||_\infty$. □

Proof of Proposition 1. Suppose that there exists $\tilde{N} \neq 0$ and $\tilde{N} \in \Omega(A^*) \cap T(B^*)$. Given that $\tilde{N} \in T(B^*)$, we can scale $\tilde{N}$ so that $\left|\left|\tilde{N}\right|\right|_2 = 1$. Thus, by Lemma 2,

$$\mu(A^*)\xi(B^*) \geq \max_{N \in T(M), ||N||_2 \leq 1} \left|\left|P_{\Omega(A^*)}(N)\right|\right|_2 \geq \left|\left|P_{\Omega(A^*)}(\tilde{N})\right|\right|_2 = 1$$

contradicting to $\mu(A^*)\xi(B^*) < 1$. The result follows. □

One important consequence of Proposition 1 is the following rank-sparsity uncertainty principle.

Theorem 3 (Rank-Sparsity Uncertainty Principle)**.** *For a matrix $M \neq 0$, we have that*

$$\xi(M) \cdot \mu(M) \geq 1.$$

Proof. Notice that $M \in \Omega(M) \cap T(M)$. By Proposition 1, we know that $\xi(M) \cdot \mu(M) < 1$, leading to a contradiction. □

2 Optimality Condition

Consider the Lagrangian function of (1) as:

$$\mathcal{L}(A, B, Q) = \gamma \left|\left|A\right|\right|_1 + \left|\left|B\right|\right|_\star + \langle Q, C - A - B\rangle.$$

From the optimality conditions of a convex program, (A^*, B^*) is a minimizer of (1) if and only if the dual matrix $Q \in \mathbb{R}^{n\times n}$ satisfies

$$Q \in \gamma\partial \left|\left|A^*\right|\right|_1 \quad \text{and} \quad Q \in \partial \left|\left|B^*\right|\right|_\star . \tag{2}$$

Based on the subdifferentials of $||\cdot||_1$ and $||\cdot||_\star$, we know that (2) is equivalent to

$$P_{\Omega(A^*)}(Q) = \gamma \text{sign}(A^*), \left|\left|P_{\Omega(A^*)}(Q)\right|\right|_\infty \leq \gamma \quad \text{and} \quad P_{T(B^*)}(Q) = UV^T, \left|\left|P_{T(B^*)^\perp}(Q)\right|\right|_2 \leq 1, \tag{3}$$

where $U, V \in \mathbb{R}^{n\times k}$ comes from $B^* = U\Sigma V^T$. (Recall that $\partial \left|\left|B^*\right|\right|_\star = \left\{UV^T + W : U^T W = WV^T = 0\right\}$.) Notice that (3) are necessary and sufficient conditions for (A^*, B^*) be **a** minimizer of (1). To ensure the uniqueness for the solution to (1), we need to tighten the conditions in (2) and (3) as the following proposition.

Proposition 4 (Uniqueness of the Optimal Solution)**.** *Suppose that $C = A^* + B^*$. Then, $\left(\hat{A}, \hat{B}\right) = (A^*, B^*)$ is the unique minimizer of (1) if the following conditions are satisfied:*

1. *$\Omega(A^*) \cap T(B^*) = \{0\}$.*
2. *There exists a dual matrix $Q \in \mathbb{R}^{n\times n}$ such that*
   - *(a) $P_{T(B^*)}(Q) = UV^T$;*
   - *(b) $P_{\Omega(A^*)}(Q) = \gamma \cdot \text{sign}(A^*)$;*
   - *(c) $\left|\left|P_{T(B^*)^\perp}(Q)\right|\right|_2 < 1$;*
   - *(d) $\left|\left|P_{\Omega(A^*)^c}(Q)\right|\right|_\infty < \gamma$.*

Proof of Proposition 4. Notice that (A^*, B^*) is an optimum by the condition 2 in Proposition 4. To avoid cluttered notation, we let $\Omega = \Omega(A^*), T = T(B^*), \Omega^c = \Omega(A^*)^c$, and $T_\perp(B^*) = T^\perp$.

Suppose that there is another feasible solution $(A^* + N_A, B^* + N_B)$ that also minimizes (1). Since $A^* + B^* = C = (A^* + N_A) + (B^* + N_B)$, we must have $N_A + N_B = 0$. For any subgradient (Q_A, Q_B) of the function $\gamma \left|\left|A\right|\right|_1 + \left|\left|B\right|\right|_\star$ at (A^*, B^*), we have that

$$\gamma \left|\left|A^* + N_A\right|\right|_1 + \left|\left|B^* + N_B\right|\right|_\star \geq \gamma \left|\left|A^*\right|\right|_1 + \left|\left|B^*\right|\right|_\star + \langle Q_A, N_A\rangle + \langle Q_B, N_B\rangle. \tag{4}$$

Since (Q_A, Q_B) is a subgradient of the function $\gamma \left|\left|A\right|\right|_1 + \left|\left|B\right|\right|_\star$ at (A^*, B^*), we must have from (3) that

- $Q_A = \gamma \cdot \text{sign}(A^*) + P_{\Omega^c}(Q_A)$ with $||P_{\Omega^c}(Q_A)||_\infty \leq \gamma$;
- $Q_B = UV^T + P_{T^\perp}(Q_B)$ with $||P_{T^\perp}(Q_B)||_2 \leq 1$.

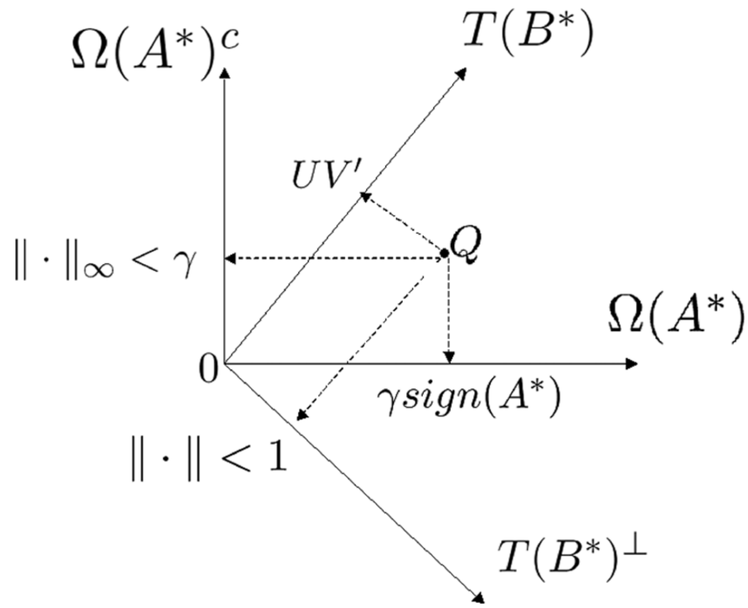

Figure 1: Geometric interpretation of optimality conditions: the existence of a dual matrix Q.

Thus, we calculate that

$$\begin{aligned}\langle Q_A, N_A\rangle &= \langle \gamma\cdot \mathrm{sign}(A^*) + P_{\Omega^c}(Q_A), N_A\rangle\\ &= \langle P_\Omega(Q) + P_{\Omega^c}(Q_A), N_A\rangle \quad \text{using (b) in Condition 2}\\ &= \langle P_{\Omega^c}(Q_A) - P_{\Omega^c}(Q), N_A\rangle + \langle Q, N_A\rangle \quad \text{by } P_\Omega(Q) = Q - P_{\Omega^c}(Q).\end{aligned}$$

Similarly, we have that

$$\begin{aligned}\langle Q_B, N_B\rangle &= \langle UV^T + P_{T^\perp}(Q_B), N_B\rangle\\ &= \langle P_T(Q) + P_{T^\perp}(Q_B), N_B\rangle \quad \text{using (a) in Condition 2}\\ &= \langle P_{T^\perp}(Q_B) - P_{T^\perp}(Q), N_B\rangle + \langle Q, N_B\rangle \quad \text{by } P_T(Q) = Q - P_{T^\perp}(Q).\end{aligned}$$

Adding the above two equalities together gives us that

$$\begin{aligned}\langle Q_A, N_A\rangle + \langle Q_B, N_B\rangle &= \langle P_{\Omega^c}(Q_A) - P_{\Omega^c}(Q), N_A\rangle + \langle Q, N_A\rangle + \langle P_{T^\perp}(Q_B) - P_{T^\perp}(Q), N_B\rangle + \langle Q, N_B\rangle\\ &= \langle P_{\Omega^c}(Q_A) - P_{\Omega^c}(Q), P_{\Omega^c}(N_A)\rangle + \langle P_{T^\perp}(Q_B) - P_{T^\perp}(Q), P_{T^\perp}(N_B)\rangle,\end{aligned} \tag{5}$$

where we use the fact that $N_A + N_B = 0$ and the projection matrices $P_{\Omega^c}, P_{T^\perp}$ are idempotent.

Given that any subgradient (Q_A, Q_B) of the function $\gamma\,||A||_1 + ||B||_\star$ at (A^*, B^*) will satisfy the above equality, we can choose (Q_A, Q_B) as follows:

- Take Q_A so that $P_{\Omega^c}(Q_A) = \gamma\cdot\mathrm{sign}(P_{\Omega^c}(N_A))$ with $||P_{\Omega^c}(Q_A)||_\infty \le \gamma$ and $\langle P_{\Omega^c}(Q_A), P_{\Omega^c}(N_A)\rangle = \gamma\,||P_{\Omega^c}(N_A)||_1$.
- Given the singular value decomposition of $P_{T^\perp}(N_B) = \tilde{U}\tilde{\Sigma}\tilde{V}^T$, we choose Q_B so that $P_{T^\perp}(Q_B) = \tilde{U}\tilde{V}^T$ with $||P_{T^\perp}(Q_B)||_2 = 1$ and $\langle P_{T^\perp}(Q_B), P_{T^\perp}(N_B)\rangle = ||P_{T^\perp}(N_B)||_\star$.

Under this choice of (Q_A, Q_B), we simplify (5) as:

$$\langle Q_A, N_A\rangle + \langle Q_B, N_B\rangle = \langle P_{\Omega^c}(Q_A) - P_{\Omega^c}(Q), P_{\Omega^c}(N_A)\rangle + \langle P_{T^\perp}(Q_B) - P_{T^\perp}(Q), P_{T^\perp}(N_B)\rangle$$

$$\geq (\gamma - ||P_{\Omega^c}(Q)||_\infty) \, ||P_{\Omega^c}(N_A)||_1 + (1 - ||P_{T^\perp}(Q)||_2) \, ||P_{T^\perp}(N_B)||_\star$$
$$> 0$$

unless $P_{\Omega^c}(N_A) = P_{T^\perp}(N_B) = 0$, where we obtain the last positivity based on (c) and (d) in Condition 2. However, if $P_{\Omega^c}(N_A) \neq 0$ or $P_{T^\perp}(N_B) \neq 0$, we know from (4) that

$$\gamma \, ||A^* + N_A||_1 + ||B^* + N_B||_\star > \gamma \, ||A^*||_1 + ||B^*||_\star \, ,$$

which violates the optimality of $(A^* + N_A, B^* + N_B)$. Now, when $P_{\Omega^c}(N_A) = P_{T^\perp}(N_B) = 0$, $P_\Omega(N_A) + P_T(N_B) = 0$ as well because of $N_A + N_B = 0$. In other words,

$$P_\Omega(N_A) = -P_T(N_B).$$

This is only possible if $P_\Omega(N_A) = P_T(N_B) = 0$ because $\Omega \cap T = \{0\}$ by Condition 1, which in turn implies that $N_A = N_B = 0$. The proof of uniqueness is completed. □

While Proposition 4 sheds light on the sufficient conditions for uniquely recovering (A^*, B^*), we now discuss the existence of an appropriate dual matrix Q entailed by Proposition 4. From Proposition 1, we already know that Condition 1 in Proposition 4 ($\Omega(A^*) \cap T(B^*) = \{0\}$) is valid when $\mu(A^*)\xi(B^*) < 1$. If we slightly strengthen the condition as $\mu(A^*)\xi(B^*) < \frac{1}{6}$, there will be a dual matrix Q satisfying the requirements in Condition 2 of Proposition 4 as well.

Theorem 5. *Given $C = A^* + B^*$ with $\mu(A^*)\xi(B^*) < \frac{1}{6}$, the unique minimizer $\left(\hat{A}, \hat{B}\right)$ of (1) will be (A^*, B^*) for the following range of γ:*

$$\gamma \in \left(\frac{\xi(B^*)}{1 - 4\mu(A^*)\xi(B^*)}, \frac{1 - 3\mu(A^*)\xi(B^*)}{\mu(A^*)} \right).$$

Specifically, $\gamma = \frac{[3\xi(B^*)]^p}{[2\mu(A^*)]^{1-p}}$ for any choice of $p \in [0, 1]$ is always inside the above range and thus guarantees exact recovery of (A^*, B^*).

The detailed proof of Theorem 5 can be found in Theorem 2 of [chandrasekaran2011rank]. The high-level idea is that we consider candidates for the dual matrix Q in the direct sum $\Omega(A^*) \oplus T(B^*)$ of the tangent spaces. Since $\mu(A^*)\xi(B^*) < \frac{1}{6}$, $\Omega(A^*) \cap T(B^*) = \{0\}$ by Proposition 1 and there exists a *unique* element $\hat{Q} \in \Omega(A^*) \oplus T(B^*)$ satisfying $P_{T(B^*)}(\hat{Q}) = UV^T$ and $P_{\Omega(A^*)}(\hat{Q}) = \gamma \cdot \mathrm{sign}(A^*)$. The proof proceeds by showing that if $\mu(A^*)\xi(B^*) < \frac{1}{6}$, then the projections of $\hat{Q}$ onto the orthogonal spaces $\Omega(A^*)^c$ and $T(B^*)^\perp$ are small, and Condition 2 of Proposition 4 is thus satisfied.

Other further reading for the course:

- Chandrasekaran, V., Recht, B., Parrilo, P. A., & Willsky, A. S. (2012). The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12, 805-849.

References

V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Sparse and low-rank matrix decompositions. *IFAC Proceedings Volumes*, 42(10):1493–1498, 2009.

V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011.

R. Hartshorne. *Algebraic geometry*, volume 52. Springer Science & Business Media, 2013.

L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.