Matrix estimation with rank constraints

Yu Zhang , Yan Chen

November 14, 2022

Outline

1. Matrix regression and applications
2. Analysis of nuclear norm regularization
   - Decomposability and subspaces
   - Restricted strong convexity and error bounds
   - Bounds under operator norm curvature
3. Matrix compressed sensing
   - noiseless observation model
   - noise observation model
4. Bounds for phase retrieval
5. Multivariate regression with low-rank constraints
6. Matrix completion
7. Additive matrix decompositions

- The analog of the Euclidean inner product on the matrix space $\mathbb{R}^{d_1 \times d_2}$ is the trace inner product

$$\langle\langle \mathbf{A}, \mathbf{B} \rangle\rangle := \operatorname{trace}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right) = \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} A_{j_1 j_2} B_{j_1 j_2}. \qquad (10.1)$$

- Frobenius norm $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} (A_{j_1 j_2})^2}$.
- In a matrix regression model, each observation takes the form $\mathbf{Z}_i = (\mathbf{X}_i, y_i)$, where $\mathbf{X}_i \in \mathbb{R}^{d_1 \times d_2}$ is a matrix of covariates, and $y_i \in \mathbb{R}$ is a response variable.
- The linear model,

$$y_i = \langle\langle \mathbf{X}_i, \Theta^* \rangle\rangle + w_i, \qquad (10.2)$$

where w_i is some type of noise variable.

- The observation operator $\mathfrak{X}_n : \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}^n$ with elements $[\mathfrak{X}_n(\mathbf{\Theta})]_i = \langle\langle \mathbf{X}_i, \mathbf{\Theta} \rangle\rangle$, and then writing this observation model in a more compact form:
$$y = \mathfrak{X}_n(\mathbf{\Theta}^*) + w, \tag{10.3}$$
where $y \in \mathbb{R}^n$ and $w \in \mathbb{R}^n$ are the vectors of response and noise variables, respectively.
- The adjoint of the observation operator: $\mathfrak{X}_n^*$, is the linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^{d_1 \times d_2}$ given by $u \mapsto \sum_{i=1}^n u_i \mathbf{X}_i$.

- There are many applications in which the regression matrix $\mathbf{\Theta}^*$ is either low-rank, or well approximated by a low-rank matrix. Thus, if we were to disregard computational costs, an appropriate estimator would be a rank-penalized form of least squares.
- However, including a rank penalty makes this a non-convex problem. This obstacle motivates replacing the rank penalty with the nuclear norm, which leads to the convex program

$$\widehat{\mathbf{\Theta}} \in \arg\min_{\mathbf{\Theta}\in\mathbb{R}^{d_1\times d_2}} \left\{ \frac{1}{2n}\|y - \mathfrak{X}_n(\mathbf{\Theta})\|_2^2 + \lambda_n |||\mathbf{\Theta}|||_{\text{nuc}} \right\}. \qquad (10.4)$$

- The nuclear norm of $\mathbf{\Theta}$ is given by the sum of its singular values:

$$|||\mathbf{\Theta}|||_{\mathrm{nuc}} = \sum_{j=1}^{d'} \sigma_j(\mathbf{\Theta}), \quad \text{where } d' = \min\{d_1, d_2\} \qquad (10.5)$$

Example 10.2 (Low-rank matrix completion)

- Matrix completion refers to the problem of estimating an unknown matrix $\mathbf{\Theta}^* \in \mathbb{R}^{d_1 \times d_2}$ based on (noisy) observations of a subset of its entries. In the linear case, we might assume that

$$\widetilde{y}_i = \Theta_{a(i),b(i)} + \frac{w_i}{\sqrt{d_1 d_2}},$$

  where w_i is some form of observation noise, and $(a(i), b(i))$ are the row and column indices of the i th observation.
- For sample index i, define the mask matrix $\mathbf{X}_i \in \mathbb{R}^{d_1 \times d_2}$, which is zero everywhere except for position $(a(i), b(i))$, where it takes the value $\sqrt{d_1 d_2}$. Then by defining the rescaled observation $y_i := \sqrt{d_1 d_2} \quad \widetilde{y}_i$, the observation model can be written in the trace regression form as

$$y_i = \langle\langle \mathbf{X}_i, \Theta^* \rangle\rangle + w_i.$$

Example 10.3 (Compressed sensing for low-rank matrices)

- Working with the linear observation model (10.3), suppose that the design matrices $\mathbf{X}_i \in \mathbb{R}^{d_1 \times d_2}$ are drawn i.i.d from a random Gaussian ensemble. In the simplest of settings, the design matrix is chosen from the standard Gaussian ensemble, meaning that each of its $D = d_1 d_2$ entries is an i.i.d. draw from the $\mathcal{N}(0,1)$ distribution.
- In this case, the random operator $\mathfrak{X}_n$ provides n random projections of the unknown matrix Θ^*-namely

$$y_i = \langle\langle \mathbf{X}_i, \mathbf{\Theta}^* \rangle\rangle \qquad \text{for } i = 1, \ldots, n.$$

In this noiseless setting, it is natural to ask how many such observations suffice to recover $\mathbf{\Theta}^*$ exactly.

Example 10.4 (Phase retrieval)

- Let $\theta^* \in \mathbb{R}^d$ be an unknown vector, and suppose that we make measurements of the form $\widetilde{y}_i = |\langle x_i, \theta^* \rangle|$ where $x_i \sim \mathcal{N}(0, \mathbf{I}_d)$ is a standard normal vector.
- Taking squares on both sides yields the equivalent observation model

$$\widetilde{y}_i^2 = (\langle x_i, \theta^* \rangle)^2 = \langle\langle x_i \otimes x_i, \theta^* \otimes \theta^* \rangle\rangle \qquad \text{for } i = 1, \ldots, n,$$

  where $\theta^* \otimes \theta^* = \theta^* (\theta^*)^{\mathrm{T}}$ is the rank-one outer product.
- By defining the scalar observation $y_i := \widetilde{y}_i^2$, as well as the matrices $\mathbf{X}_i := x_i \otimes x_i$ and $\mathbf{\Theta}^* := \theta^* \otimes \theta^*$, we obtain an equivalent version of the noiseless phase retrieval problem-namely, to find a rank-one solution to the set of matrix-linear equations $y_i = \langle\langle \mathbf{X}_i, \mathbf{\Theta}^* \rangle\rangle$ for $i = 1, \ldots, n$.

For any given matrix $\mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2}$, we let rowspan $(\mathbf{\Theta}) \subseteq \mathbb{R}^{d_2}$ and colspan $(\mathbf{\Theta}) \subseteq \mathbb{R}^{d_1}$ denote its row space and column space, respectively. For a given positive integer $r \leq d' := \min\{d_1, d_2\}$, let $\mathbb{U}$ and $\mathbb{V}$ denote r-dimensional subspaces of vectors. Define the two subspaces of matrices

- $$\mathbb{M}(\mathbb{U}, \mathbb{V}) := \left\{ \mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2} \mid \mathrm{rowspan}(\mathbf{\Theta}) \subseteq \mathbb{V}, \mathrm{colspan}(\mathbf{\Theta}) \subseteq \mathbb{U} \right\} \tag{10.12a}$$

- $$\bar{\mathbb{M}}^{\perp}(\mathbb{U}, \mathbb{V}) := \left\{ \mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2} \mid \mathrm{rowspan}(\mathbf{\Theta}) \subseteq \mathbb{V}^{\perp}, \quad \mathrm{colspan}(\mathbf{\Theta}) \subseteq \mathbb{U}^{\perp} \right\} \tag{10.12b}$$

Here $\mathbb{U}^{\perp}$ and $\mathbb{V}^{\perp}$ denote the subspaces orthogonal to $\mathbb{U}$ and $\mathbb{V}$, respectively. On the other hand, equation $(10.12b)$ defines the subspace $\bar{\mathbb{M}}(\mathbb{U}, \mathbb{V})$ implicitly, via taking the orthogonal complement.

- Let $\mathbf{U} \in \mathbb{R}^{d_1 \times d'}$ and $\mathbf{V} \in \mathbb{R}^{d_2 \times d'}$ be a pair of orthonormal matrices. These matrices can be used to define r-dimensional spaces: namely, let $\mathbb{U}$ be the span of the first r columns of $\mathbf{U}$, and similarly, let $\mathbb{V}$ be the span of the first r columns of $\mathbf{V}$.
- In practice, these subspaces correspond (respectively) to the spaces spanned by the top r left and right singular vectors of the target matrix $\mathbf{\Theta}^*$.

With these choices, any pair of matrices $\mathbf{A} \in \mathbb{M}(\mathbb{U}, \mathbb{V})$ and $\mathbf{B} \in \bar{\mathbb{M}}^{\perp}(\mathbb{U}, \mathbb{V})$ can be represented in the form

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{0}_{r\times(d'-r)} \\ \mathbf{0}_{(d'-r)\times r} & \mathbf{0}_{(d'-r)\times(d'-r)} \end{bmatrix} \mathbf{V}^{\mathrm{T}}, \mathbf{B} = \mathbf{U} \begin{bmatrix} \mathbf{0}_{r\times r} & \mathbf{0}_{r\times(d'-r)} \\ \mathbf{0}_{(d'-r)\times r} & \mathbf{\Gamma}_{22} \end{bmatrix} \mathbf{V}^{\mathrm{T}}$$

where $\mathbf{\Gamma}_{11} \in \mathbb{R}^{r\times r}$ and $\mathbf{\Gamma}_{22} \in \mathbb{R}^{(d'-r)\times(d'-r)}$ are arbitrary matrices. Since the trace inner product defines orthogonality, any member $\overline{\mathbf{A}}$ of $\bar{M}(\mathbb{U}, \mathbb{V})$ must take the form

$$\overline{\mathbf{A}} = \mathbf{U} \begin{bmatrix} \overline{\mathbf{\Gamma}}_{11} & \overline{\mathbf{\Gamma}}_{12} \\ \overline{\mathbf{\Gamma}}_{21} & \mathbf{0} \end{bmatrix} \mathbf{V}^{\mathrm{T}},$$

where all three matrices $\overline{\mathbf{\Gamma}}_{11} \in \mathbb{R}^{r\times r}, \overline{\mathbf{\Gamma}}_{12} \in \mathbb{R}^{r\times(d'-r)}$ and $\overline{\mathbf{\Gamma}}_{21} \in \mathbb{R}^{(d'-r)\times r}$ are arbitrary.

- In this way, we see explicitly that $\bar{\mathbb{M}}$ is a strict superset of $\mathbb{M}$ whenever $r < d'$. Whereas any matrix in $\mathbb{M}$ has rank at most r, the representation shows that any matrix in $\bar{\mathbb{M}}$ has rank at most $2r$.
- The preceding discussion also demonstrates the decomposability of the nuclear norm. For an arbitrary pair of matrices $\mathbf{A} \in \mathbb{M}$ and $\mathbf{B} \in \bar{\mathbb{M}}^{\perp}$, we have

$$\begin{aligned}
|||\mathbf{A} + \mathbf{B}|||_{\mathsf{nuc}} &\overset{\text{(i)}}{=} ||| \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_{22} \end{bmatrix} |||_{\mathrm{nuc}} \\
&= ||| \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} |||_{\mathsf{nuc}} + ||| \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_{22} \end{bmatrix} |||_{\mathrm{nuc}} \\
&\overset{\text{(ii)}}{=} |||\mathbf{A}|||_{\mathsf{nuc}} + \|\mathbf{B}\|_{\mathrm{nuc}},
\end{aligned}$$

where steps (i) and (ii) use the invariance of the nuclear norm to orthogonal transformations.

- Consider an M-estimator of the form

$$\widehat{\boldsymbol{\Theta}} \in \arg \min_{\boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2}} \left\{ \mathcal{L}_n(\boldsymbol{\Theta}) + \lambda_n |||\boldsymbol{\Theta}|||_{\mathrm{nuc}} \right\},$$

where $\mathcal{L}_n$ is some convex and differentiable cost function.

- Then for any choice of regularization parameter $\lambda_n \geq 2|||\nabla \mathcal{L}_n (\boldsymbol{\Theta}^*)|||_2$, the error matrix $\widehat{\boldsymbol{\Delta}} = \widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}^*$ must satisfy the cone-like constraint

$$|||\widehat{\boldsymbol{\Delta}}_{\bar{\mathbb{M}}^\perp}|||_{\mathrm{nuc}} \leq 3|||\widehat{\boldsymbol{\Delta}}_{\bar{\mathbb{M}}}|||_{\mathrm{nuc}} + 4|||\Theta^*_{\mathbb{M}^\perp}|||_{\mathrm{nuc}} \tag{10.15}$$

where $\mathbb{M} = \mathbb{M}(\mathbb{U}^r, \mathbb{V}^r)$ and $\bar{\mathbb{M}} = \bar{\mathbb{M}}(\mathbb{U}^r, \mathbb{V}^r)$.

(

$$\mathbb{G}(\lambda_n) := \left\{ \Phi^* \left(\nabla \mathcal{L}_n (\theta^*) \right) \leq \frac{\lambda_n}{2} \right\}.$$

Proposition 9.13

Let $\mathcal{L}_n : \Omega \to \mathbb{R}$ be a convex function, let the regularizer $\Phi : \Omega \to [0, \infty)$ be a norm, and consider a subspace pair $\left(\mathbb{M}, \bar{\mathbb{M}}^{\perp}\right)$ over which Φ is decomposable. Then conditioned on the event $\mathbb{G}(\lambda_n)$, the error $\widehat{\Delta} = \widehat{\theta} - \theta^*$ belongs to the set

$$\mathbb{C}_{\theta'}\left(\mathbb{M}, \bar{\mathbb{M}}^{\perp}\right) := \left\{ \Delta \in \Omega \mid \Phi\left(\Delta_{\bar{\mathbb{M}}^{\perp}}\right) \leq 3\Phi\left(\Delta_{\overline{\mathbb{M}}}\right) + 4\Phi\left(\theta^*_{\mathbb{M}^{\perp}}\right) \right\}.$$

)

- Given observations $(y, \mathfrak{X}_n)$ from the matrix regression model (10.3), consider the estimator

$$\widehat{\mathbf{\Theta}} \in \arg \min_{\mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2}} \left\{ \frac{1}{2n} \|y - \mathfrak{X}_n(\mathbf{\Theta})\|_2^2 + \lambda_n |||\mathbf{\Theta}|||_{\text{nuc}} \right\}, \qquad (10.16)$$

where $\lambda_n > 0$ is a user-defined regularization parameter.
- The nuclear norm is a decomposable regularizer and the least-squares cost is convex, and so given a suitable choice of λ_n, the error matrix $\widehat{\mathbf{\Delta}} := \widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*$ must satisfy the cone-like constraint (10.15).

The restricted strong convexity of the loss function. For this least-squares cost, we show the random operator $\mathfrak{X}_n$ satisfies a uniform lower bound of the form

$$\frac{\|\mathfrak{X}_n(\mathbf{\Delta})\|_2^2}{2n} \geq \frac{\kappa}{2}|||\mathbf{\Delta}|||_{\mathrm{F}}^2 - c_0\frac{(d_1+d_2)}{n}|||\mathbf{\Delta}|||_{\mathrm{nuc}}^2, \quad \text{for all } \mathbf{\Delta} \in \mathbb{R}^{d_1 \times d_2} \tag{10.17}$$

with high probability. Here the quantity $\kappa > 0$ is a curvature constant, and c_0 is another universal constant of secondary importance.

Proposition 10.6

Suppose that the observation operator $\mathfrak{X}_n$ satisfies the restricted strong convexity condition (10.17) with parameter $\kappa > 0$. Then conditioned on the event $\mathbb{G}(\lambda_n) = \left\{ \left\| \frac{1}{n} \sum_{i=1}^{n} w_i \mathbf{X}_i \right\|_2 \leq \frac{\lambda_n}{2} \right\}$, any optimal solution to nuclear norm regularized least squares (10.16) satisfies the bound

$$
\begin{aligned}
|||\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*|||_{\mathrm{F}}^2 \leq \frac{9}{2} \frac{\lambda_n^2}{\kappa^2} r + \frac{1}{\kappa} \{ 2\lambda_n \sum_{j=r+1}^{d'} \sigma_j(\mathbf{\Theta}^*) \} \\
+ \frac{1}{\kappa} \{ \frac{32 c_0 (d_1 + d_2)}{n} \left[\sum_{j=r+1}^{d'} \sigma_j(\mathbf{\Theta}^*) \right]^2 \}
\end{aligned} \tag{10.18}
$$

valid for any $r \in \{1, \ldots, d'\}$such that$r \leq \frac{\kappa n}{128 c_0 (d_1 + d_2)}$

Remark:

- It is splitting into estimation and approximation error, parameterized by the choice of r. Note that the choice of r can be optimized so as to obtain the tightest possible bound.
- Suppose that $\operatorname{rank}(\boldsymbol{\Theta}^*) < d'$ and moreover that $n > 128\frac{c_0}{k}\operatorname{rank}(\boldsymbol{\Theta}^*)(d_1 + d_2)$. We then may apply the bound (10.18) with $r = \operatorname{rank}(\boldsymbol{\Theta}^*)$. Proposition 10.6 implies the upper bound

$$|||\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}^*|||_{\mathrm{F}}^2 \leq \frac{9}{2}\frac{\lambda_n^2}{\kappa^2}\operatorname{rank}(\boldsymbol{\Theta}^*).$$

- Theorem 9.19 (Bounds for general models)
  Under conditions (A1) and (A2), consider the regularized M-estimator (9.3) conditioned on the event $\mathbb{G}(\lambda_n)$,
  (a) Any optimal solution satisfies the bound

$$\Phi\left(\widehat{\theta}-\theta^*\right) \leq 4\left\{\Psi(\bar{\mathbb{M}})\left\|\widehat{\theta}-\theta^*\right\| + \Phi\left(\theta^*_{\mathbb{M}^\perp}\right)\right\}.$$

  (b) For any subspace pair $\left(\overline{\mathbb{M}}, \mathbb{M}^\perp\right)$ such that $\tau_n^2\Psi^2(\overline{\mathbb{M}}) \leq \frac{\kappa}{64}$ and $\varepsilon_n\left(\overline{\mathbb{M}}, \mathbb{M}^\perp\right) \leq R$, we have

$$\left\|\widehat{\theta}-\theta^*\right\|^2 \leq \varepsilon_n^2\left(\overline{\mathbb{M}}, \mathbb{M}^\perp\right).$$

$$\varepsilon_n^2\left(\overline{\mathbb{M}}, \mathbb{M}^\perp\right) := \underbrace{9\frac{\lambda_n^2}{\kappa^2}\Psi^2(\overline{\mathbb{M}})}_{\text{estimation error}} + \underbrace{\frac{8}{\kappa}\left\{\lambda_n\Phi\left(\theta^*_{\mathbb{M}^\perp}\right) + 16\tau_n^2\Phi^2\left(\theta^*_{\mathbb{M}^\perp}\right)\right\}}_{\text{approximation error}},$$

Proof:

- For each $r \in \{1, \ldots, d'\}$, let $(\mathbb{U}^r, \mathbb{V}^r)$ be the subspaces spanned by the top r left and right singular vectors of Θ^*: $\mathbb{M}(\mathbb{U}^r, \mathbb{V}^r)$ and $\bar{\mathbb{M}}^\perp(\mathbb{U}^r, \mathbb{V}^r)$. As shown previously, the nuclear norm is decomposable with respect to any such subspace pair.
- The dual norm to the nuclear norm is the ℓ_2-operator norm. For the least-squares cost function, $\nabla \mathcal{L}_n(\boldsymbol{\Theta}^*) = \frac{1}{n} \sum_{i=1}^n w_i \mathbf{X}_i$. So the "good" event$\mathbb{G}(\lambda_n)$ is given.

- The assumption (10.17) is a form of restricted strong convexity with tolerance parameter $\tau_n^2 = c_0 \frac{d_1+d_2}{n}$.
- It only remains to verify the condition $\tau_n^2 \Psi^2(\bar{\mathbb{M}}) \leq \frac{k}{64}$.
The representation (10.14) reveals that any matrix $\Theta \in \bar{M}\left(\mathbb{U}^r, \mathbb{V}^r\right)$ has rank at most $2r$, and hence

$$\Psi\left(\overline{\mathbb{M}}\left(\mathbb{U}^r, \mathbb{V}^r\right)\right) := \sup_{\boldsymbol{\Theta} \in \bar{\mathbb{M}}(\mathbb{U}^r, \mathbb{V}^r) \setminus \{0\}} \frac{|||\boldsymbol{\Theta}|||_{\mathrm{nuc}}}{|||\boldsymbol{\Theta}|||_{\mathrm{F}}} \leq \sqrt{2r}.$$

- Consequently, the final condition of Theorem 9.19 holds whenever the target rank r is bounded as in the statement of Proposition 10.6.

For a given regularizer Φ, provides a bound on the estimation error in terms of the dual norm Φ^*.

- The dual to the nuclear norm is the ℓ_2-operator norm or spectral norm. For the least-squares cost function, the gradient is given by

$$\nabla \mathcal{L}_n(\mathbf{\Theta}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i^{\mathrm{T}} \left(y_i - \langle\langle \mathbf{X}_i, \mathbf{\Theta}\rangle\rangle\right) = \frac{1}{n}\mathfrak{X}_n^* \left(y - \mathfrak{X}_n(\mathbf{\Theta})\right),$$

  where $\mathfrak{X}_n^* : \mathbb{R}^n \to \mathbb{R}^{d_1 \times d_2}$ is the adjoint operator.
- The Φ^*-curvature condition from Definition 9.22 takes the form

$$|||\frac{1}{n}\mathfrak{X}_n^*\mathfrak{X}_n(\mathbf{\Delta})|||_2 \geq \kappa |||\mathbf{\Delta}|||_2 - \tau_n |||\mathbf{\Delta}|||_{\text{nuc}} \qquad \text{for all } \mathbf{\Delta} \in \mathbb{R}^{d_1 \times d_2},$$

  where $\kappa > 0$ is the curvature parameter, and $\tau_n \geq 0$ is the tolerance parameter.

Proposition 10.7

Suppose that the observation operator $\mathfrak{X}_n$ satisfies the curvature condition (10.20) with parameter $\kappa > 0$, and consider a matrix $\mathbf{\Theta}^*$ with $\operatorname{rank}(\mathbf{\Theta}^*) < \frac{\kappa}{64\tau_n}$. Then, conditioned on the event $\mathbb{G}(\lambda_n) = \left\{ \left\| \frac{1}{n}\mathfrak{X}_n^*(w) \right\|_2 \leq \frac{\lambda_n}{2} \right\}$, any optimal solution to the M-estimator (10.16) satisfies the bound

$$\left\| \widehat{\mathbf{\Theta}} - \mathbf{\Theta}^* \right\|_2 \leq 3\sqrt{2}\frac{\lambda_n}{\kappa}$$

Recall:

Theorem 9.24 :Given a target parameter $\theta^* \in \mathbb{M}$, consider the regularized M-estimator (9.3) under conditions $(A1')$ and $(A2)$, and suppose that $\tau_n \Psi^2(\bar{\mathbb{M}}) < \frac{k}{32}$. Conditioned on the event $\mathbb{G}(\lambda_n) \cap \left\{\Phi^*\left(\widehat{\theta} - \theta^*\right) \le R\right\}$, any optimal solution $\widehat{\theta}$ satisfies the bound

$$\Phi^*\left(\widehat{\theta} - \theta^*\right) \le 3\frac{\lambda_n}{\kappa}.$$

Proof:

In order to apply Theorem 9.24, the only remaining condition to verify is the inequality $\tau_n \Psi^2(\bar{\mathbb{M}}) < \frac{K}{32}$. We have previously calculated that $\Psi^2(\bar{\mathbb{M}}) \le 2r$, so that the stated upper bound on r ensures that this inequality holds.

- $\mathbf{X}_i$ is drawn from the $\boldsymbol{\Sigma}$-Gaussian ensemble: One might draw random observation matrices $\mathbf{X}_i$ with dependent entries, for instance with $\operatorname{vec}(\mathbf{X}_i) \sim \mathcal{N}(0, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} \in \mathbb{R}^{(d_1 d_2) \times (d_1 d_2)}$ is the covariance matrix.
- $$\rho^2(\boldsymbol{\Sigma}) := \sup_{\|u\|_2 = \|v\|_2 = 1} \operatorname{var}\left(\left\langle\left\langle \mathbf{X}, uv^{\mathrm{T}} \right\rangle\right\rangle\right).$$

  Note that $\rho^2(\mathbf{I}_d) = 1$ for the special case of the identity ensemble.

Theorem 10.8

Given n i.i.d. draws $\{X_i\}_{i=1}^n$ of random matrices from the $\mathbf{\Sigma}$-Gaussian ensemble, there are positive constants $c_1 < 1 < c_2$ such that

$$\frac{\|\mathfrak{X}_n(\mathbf{\Delta})\|_2^2}{n} \geq c_1 \|\sqrt{\mathbf{\Sigma}}\,\mathrm{vec}(\mathbf{\Delta})\|_2^2 - c_2 \rho^2(\mathbf{\Sigma}) \left\{ \frac{d_1 + d_2}{n} \right\} \|\mathbf{\Delta}\|_{\mathrm{nuc}}^2 \quad \forall \mathbf{\Delta} \in \mathbb{R}^{d_1 \times d_2}$$

with probability at least $1 - \frac{e^{-\frac{n}{32}}}{1 - e^{-\frac{n}{32}}}$.

This result can be understood as a variant of Theorem 7.16, which established a similar result for the case of sparse vectors and the ℓ_1-norm. As with this earlier theorem, Theorem 10.8 can be proved using the Gordon-Slepian comparison lemma for Gaussian processes. In Exercise 10.6, we work through a proof of a slightly simpler form of the bound.

Consider following convex program:

$$\min_{\boldsymbol{\Theta}\in\mathbb{R}^{d_1\times d_2}} |||\boldsymbol{\Theta}|||_{\mathrm{nuc}} \quad \text{such that } \langle\langle \mathbf{X}_i, \boldsymbol{\Theta}\rangle\rangle = y_i \text{ for all } i = 1, \ldots, n \quad (10.23)$$

That is, we search over the space of matrices that match the observations perfectly to find the solution with minimal nuclear norm.

Corollary 10.9

Given $n > 16\frac{c_2}{c_1}\frac{\rho^2(\boldsymbol{\Sigma})}{\gamma_{\min}(\boldsymbol{\Sigma})}r\,(d_1 + d_2)$ i.i.d. samples from the $\boldsymbol{\Sigma}$-ensemble, the estimator (10.23) recovers the rank-r matrix $\boldsymbol{\Theta}^*$ exactly-i.e., it has a unique solution $\widehat{\boldsymbol{\Theta}} = \boldsymbol{\Theta}^*$ with probability at least $1 - \frac{e^{-\frac{n}{32}}}{1-e^{-\frac{n}{32}}}$.

Proof:
Since $\widehat{\Theta}$ and Θ^* are optimal and feasible, respectively, for the program (10.23), we have $|||\widehat{\Theta}|||_{\text{nuc}} \leq |||\Theta^*|||_{\text{nuc}} = |||\Theta^*_{\mathbb{M}}|||_{\text{nuc}}$. Introducing the error matrix $\widehat{\Delta} = \widehat{\Theta} - \Theta^*$, we have

$$|||\widehat{\Theta}\|_{\mathrm{nuc}} = |||\Theta^* + \widehat{\Delta}|||_{\mathrm{nuc}} = |||\Theta^*_{\mathbb{M}} + \widehat{\Delta}_{\bar{\mathbb{M}}^\perp} + \widehat{\Delta}_{\bar{\mathbb{M}}}|||_{\mathrm{nuc}}$$
$$\geq |||\Theta^*_{\mathbb{M}} + \widehat{\Delta}_{\bar{\mathbb{M}}^\perp}|||_{\mathrm{nuc}} - |||\widehat{\Delta}_{\bar{\mathbb{M}}}|||_{\mathrm{nuc}}$$

Applying decomposability this yields
$|||\Theta^*_{\mathbb{M}} + \widehat{\Delta}_{\bar{\mathbb{M}}^\perp}|||_{\mathrm{nuc}} = |||\Theta^*_{\mathbb{M}}|||_{\mathrm{nuc}} + |||\widehat{\Delta}_{\bar{\mathbb{M}}^\perp}|||_{\mathrm{nuc}}$
Combining the pieces, we find that $|||\widehat{\Delta}_{\bar{\mathbb{M}}^\perp}|||_{\mathrm{nuc}} \leq |||\widehat{\Delta}_{\bar{\mathbb{M}}}|||_{\mathrm{nuc}}$
From the representation (10.14), any matrix in $\bar{M}$ has rank at most $2r$, whence

$$|||\widehat{\Delta}|||_{\text{nuc}} \leq 2|||\widehat{\Delta}_{\bar{\mathbb{M}}}|||_{\mathrm{nuc}} \leq 2\sqrt{2r}|||\widehat{\Delta}|||_{\mathrm{F}} \tag{10.24}$$

Now let us condition on the event that the lower bound (10.22) holds. When applied to $\widehat{\Delta}$, and coupled with the inequality (10.24), we find that

$$\begin{aligned} \frac{\left\| \mathfrak{X}_n(\widehat{\mathbf{\Delta}}) \right\|_2^2}{n} &\geq \left\{ c_1 \gamma_{\min}(\mathbf{\Sigma}) - 8 c_2 \rho^2(\mathbf{\Sigma}) \frac{r\,(d_1 + d_2)}{n} \right\} |||\widehat{\mathbf{\Delta}}|||_{\mathrm{F}}^2 \\ &\geq \frac{c_1}{2} \gamma_{\min}(\mathbf{\Sigma}) |||\widehat{\mathbf{\Delta}}|||_{\mathrm{F}}^2, \end{aligned} \tag{1}$$

where the final inequality follows by applying the given lower bound on n, and performing some algebra. But since both $\widehat{\mathbf{\Theta}}$ and $\mathbf{\Theta}^*$ are feasible for the convex program (10.23), we have shown that $0 = \frac{\left\| \mathfrak{X}_n(\widehat{\mathbf{\Delta}}) \right\|_2^2}{n} \geq \frac{c_1}{2} \gamma_{\min}(\mathbf{\Sigma}) |||\widehat{\mathbf{\Delta}}|||_{\mathrm{F}}^2$, which implies that $\widehat{\mathbf{\Delta}} = 0$ as claimed.

Consider the estimator

$$\widehat{\boldsymbol{\Theta}} \in \arg \min_{\boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2}} \left\{ \frac{1}{2n} \|y - \mathfrak{X}_n(\boldsymbol{\Theta})\|_2^2 + \lambda_n |||\boldsymbol{\Theta}|||_{\text{nuc}} \right\}, \tag{10.16}$$

based on noisy observations of the form $y_i = \langle\langle \mathbf{X}_i, \boldsymbol{\Theta}^* \rangle\rangle + w_i$.

Corollary 10.10

Consider $n > 64 \frac{c_2}{c_1} \frac{\rho^2(\boldsymbol{\Sigma})}{\gamma_{\min}(\boldsymbol{\Sigma})} r\,(d_1 + d_2)$ i.i.d. samples $(y_i, \mathbf{X}_i)$ from the linear matrix regression model, where each $\mathbf{X}_i$ is drawn from the $\boldsymbol{\Sigma}$-Gaussian ensemble. Then any optimal solution to the program (10.16) with $\lambda_n = 10\sigma\rho(\boldsymbol{\Sigma}) \left(\sqrt{\frac{d_1+d_2}{n}} + \delta \right)$ satisfies the bound

$$|||\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}^*|||_{\mathrm{F}}^2 \leq 125 \frac{\sigma^2 \rho^2(\boldsymbol{\Sigma})}{c_1^2 \gamma_{\min}^2(\boldsymbol{\Sigma})} \left\{ \frac{r\,(d_1 + d_2)}{n} + r\delta^2 \right\} \tag{10.25}$$

with probability at least $1 - 2e^{-2n\delta^2}$.

Proof:

- We prove the bound (10.25) via an application of Proposition 10.6, in particular in the form of the bound (10.19). Theorem 10.8 shows that the RSC condition holds with $\kappa = c_1 \gamma_{\min}(\Sigma)$ and $c_0 = \frac{c_2 \rho^2(\mathbf{\Sigma})}{2}$, so that the stated lower bound on the sample size ensures that Proposition 10.6 can be applied with $r = \operatorname{rank}(\Theta^*)$.
- It remains to verify that the event $\mathbb{G}(\lambda_n) = \left\{ |||\frac{1}{n}\sum_{i=1}^n w_i \mathbf{X}_i|||_2 \leq \frac{\lambda_n}{2} \right\}$ holds with high probability. Introduce the shorthand $\mathbf{Q} = \frac{1}{n}\sum_{i=1}^n w_i \mathbf{X}_i$, and define the event $\mathcal{E} = \left\{ \frac{\|w\|_2^2}{n} \leq 2\sigma^2 \right\}$. We then have

$$\mathbb{P}\left[|||\mathbf{Q}|||_2 \geq \frac{\lambda_n}{2} \right] \leq \mathbb{P}[\mathcal{E}^c] + \mathbb{P}\left[|||\mathbf{Q}|||_2 \geq \frac{\lambda_n}{2} \mid \mathcal{E} \right].$$

  Since the noise variables $\{w_i\}_{i=1}^n$ are i.i.d., each zero-mean and sub-Gaussian with parameter σ, we have $\mathbb{P}[\mathcal{E}^c] \leq e^{-n/8}$.

Let $\left\{u^1, \ldots, u^M\right\}$ and $\left\{v^1, \ldots, v^N\right\}$ be 1/4-covers in Euclidean norm of the spheres $\mathbb{S}^{d_1-1}$ and $\mathbb{S}^{d_2-1}$, respectively. By Lemma 5.7, we can find such covers with $M \leq 9^{d_1}$ and $N \leq 9^{d_2}$ elements respectively. For any $v \in \mathbb{S}^{d_2-1}$, we can write $v = v^\ell + \Delta$ for some vector Δ with ℓ_2 at most 1/4, and hence

$$|||\mathbf{Q}|||_2 = \sup_{v \in \mathbb{S}_{d_2-1}} \|\mathbf{Q}v\|_2 \leq \frac{1}{4}|||\mathbf{Q}|||_2 + \max_{\ell=1,\ldots,N} \left\|\mathbf{Q}v^\ell\right\|_2 .$$

A similar argument involving the cover of $\mathbb{S}^{d_1-1}$ yields $\left\|\mathbf{Q}v^\ell\right\|_2 \leq \frac{1}{4}|||\mathbf{Q}|||_2 + \max_{j=1,\ldots,M} \left\langle u^j, \mathbf{Q}v^\ell \right\rangle$. Thus, we have established that

$$\|\mathbf{Q}\|_2 \leq 2 \max_{j=1,\ldots,M} \max_{\ell=1,\ldots,N} \left|Z^{j,\ell}\right| \qquad \text{where } Z^{j,\ell} = \left\langle u^j, \mathbf{Q}v^\ell \right\rangle$$

Fix some index pair (j, ℓ) : we can then write $Z^{j,\ell} = \frac{1}{n} \sum_{i=1}^{n} w_i Y_i^{j,\ell}$ where $Y_i^{j,\ell} = \left\langle u^j, \mathbf{X}_i v^\ell \right\rangle$.
Note that each variable $Y_i^{j,\ell}$ is zero-mean Gaussian with variance at most $\rho^2(\mathbf{\Sigma})$.
Consequently, the variable $Z^{j,\ell}$ is zero-mean Gaussian with variance at most $\frac{2\sigma^2\rho^2(\mathbf{\Sigma})}{n}$, where we have used the conditioning on event $\mathcal{E}$.

Putting together the pieces, we conclude that

$$
\begin{aligned}
\mathbb{P}\left[|||\frac{1}{n}\sum_{i=1}^{n} w_i \mathbf{X}_i|||_2 \geq \frac{\lambda_n}{2} \mid \mathcal{E} \right] &\leq \sum_{j=1}^{M}\sum_{\ell=1}^{N} \mathbb{P}\left[\left| Z^{j,\ell} \right| \geq \frac{\lambda_n}{4} \right] \\
&\leq 2e^{-\frac{n\lambda_n^2}{32\sigma^2\rho^2(\mathbf{\Sigma})} + \log M + \log N} \\
&\leq 2e^{-\frac{n\lambda_n^2}{32\sigma^2\rho^2(\mathbf{\Sigma})} + (d_1+d_2)\log 9}
\end{aligned}
$$

Setting $\lambda_n = 10\sigma\rho(\mathbf{\Sigma})\left(\sqrt{\frac{(d_1+d_2)}{n}} + \delta\right)$ we find that $\mathbb{P}\left[\left\| \frac{1}{n}\sum_{i=1}^{n} w_i \mathbf{X}_i \right\|_2 \geq \frac{\lambda_n}{2} \right] \leq 2e^{-2n\delta^2}$ as claimed.

- We need to find a rank-one solution to the set of matrix-linear equations $y_i = \langle\langle \mathbf{X}_i, \quad \mathbf{\Theta}^* \rangle\rangle$ for $i = 1, \ldots, n$.
-
$$\widehat{\mathbf{\Theta}} \in \arg \min_{\mathbf{\Theta} \in \mathcal{S}_+^{d \times d}} \operatorname{trace}(\mathbf{\Theta}) \tag{10.29}$$
such that $\widetilde{y}_i^2 = \langle\langle \mathbf{\Theta}, x_i \otimes x_i \rangle\rangle \quad$ for all $i = 1, \ldots, n$.

Theorem 10.12(Restricted nullspace/eigenvalues for phase retrieval)

For each $i = 1, \ldots, n$, consider random matrices of the form $\mathbf{X}_i = x_i \otimes x_i$ for i.i.d. $\mathcal{N}(0, \mathbf{I}_d)$ vectors. Then there are universal constants (c_0, c_1, c_2) such that for any $\rho > 0$, a sample size $n > c_0 \rho d$ suffices to ensure that

$$\frac{1}{n}\sum_{i=1}^{n} \langle\langle \mathbf{X}_i, \mathbf{\Theta} \rangle\rangle^2 \geq \frac{1}{2}|||\mathbf{\Theta}|||_{\mathrm{F}}^2 \quad \text{for all matrices such that } |||\mathbf{\Theta}|||_{\mathrm{F}}^2 \leq \rho |||\mathbf{\Theta}|||_{\mathrm{nuc}}^2 \tag{10.30}$$

with probability at least $1 - c_1 e^{-c_2 n}$.

Exercise 10.9 (Phase retrieval with Gaussian masks)
Recall the real-valued phase retrieval problem, based on the functions $f_{\boldsymbol{\Theta}}(\mathbf{X}) = \langle\langle \mathbf{X}, \boldsymbol{\Theta} \rangle\rangle$, for a random matrix $\mathbf{X} = x \otimes x$ with $x \sim \mathcal{N}(0, \mathbf{I}_n)$
(a) Letting $\boldsymbol{\Theta} = \mathbf{U}^{\mathrm{T}} \mathbf{D} \mathbf{U}$ denote the singular value decomposition of $\boldsymbol{\Theta}$, explain why the random variables $f_{\boldsymbol{\Theta}}(\mathbf{X})$ and $f_{\mathbf{D}}(\mathbf{X})$ have the same distributions.
(b) Prove that

$$\mathbb{E}\left[f_{\Theta}^2(\mathbf{X})\right] = |||\boldsymbol{\Theta}|||_F^2 + 2(\operatorname{trace}(\boldsymbol{\Theta}))^2.$$

$\mathrm{E}(X) = \mu, \operatorname{Cov}(X) = \Sigma$, then

$$\mathrm{E}\left(X' A X\right) = \mu' A \mu + \operatorname{tr}(A\Sigma)$$

$$Var\left(X' A X\right) = \left(m_4 - 3\sigma^4\right) a' a + 2\sigma^4 \operatorname{tr}\left(A^2\right) + 4\sigma^2 \mu' A^2 \mu + 4 m_3 \mu' A a$$

Corollary 10.13

Given $n > 2c_0$ d samples, the SDP(10.29) has the unique optimal solution $\widehat{\mathbf{\Theta}} = \mathbf{\Theta}^*$ with probability at least $1 - c_1 e^{-c_2 n}$.

Proof: Since $\widehat{\mathbf{\Theta}}$ and $\mathbf{\Theta}^*$ are optimal and feasible (respectively) for the convex program (10.29), we are guaranteed that trace$(\widehat{\mathbf{\Theta}}) \leq$ trace $(\mathbf{\Theta}^*)$. Since both matrices are positive semidefinite, this trace constraint is equivalent to $|||\widehat{\mathbf{\Theta}}|||_{\mathrm{nuc}} \leq |||\mathbf{\Theta}^*|||_{\mathrm{nuc}}$.

In conjunction with the rank-one nature of $\mathbf{\Theta}^*$ and the decomposability of the nuclear norm, implies that the error matrix $\widehat{\Delta} = \widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*$ satisfies the cone constraint $|||\widehat{\Delta}|||_{\text{nuc}} \leq \sqrt{2}|||\widehat{\Delta}|||_{\mathrm{F}}$.
Consequently, we can apply Theorem 10.12 with $\rho = 2$ to conclude that

$$0 = \frac{1}{n} \sum_{i=1}^{n} \left\langle \left\langle \mathbf{X}_i, \widehat{\Delta} \right\rangle \right\rangle \geq \frac{1}{2} \|\widehat{\Delta}\|_2^2,$$

Outline

Model: Multivariate regression with low-rank constraints

In the case of linear prediction, any such mapping can be parameterized by a matrix $\mathbf{\Theta}^* \in \mathbb{R}^{p \times T}$. A collection of n observations can be specified by the model

$$\mathbf{Y} = \mathbf{Z}\mathbf{\Theta}^* + \mathbf{W}, \tag{2}$$

where $(\mathbf{Y}, \mathbf{Z}) \in \mathbb{R}^{n \times T} \times \mathbb{R}^{n \times p}$ are observed, and $\mathbf{W} \in \mathbb{R}^{n \times T}$ is a matrix of noise variables.

- The least-squares cost function is $\mathcal{L}_n(\mathbf{\Theta}) = \frac{1}{2n}\|\mathbf{Y} - \mathbf{Z}\mathbf{\Theta}\|_{\mathrm{F}}^2$.
- It is applicable to the case of fixed design and so involves the minimum and maximum eigenvalues of the sample covariance matrix $\widehat{\mathbf{\Sigma}} := \frac{\mathbf{Z}^{\mathrm{T}}\mathbf{Z}}{n}$.

Consider the estimator with nuclear norm

$$\widehat{\mathbf{\Theta}} \in \arg \min_{\mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2}} \left\{ \frac{1}{2n} \|y - \mathbf{Z}\mathbf{\Theta}\|_2^2 + \lambda_n \|\mathbf{\Theta}\|_{\text{nuc}} \right\}, \tag{3}$$

where $\lambda_n > 0$ is a user-defined regularization parameter.

Corollary 10.14

Consider the observation model (2) in which $\mathbf{\Theta}^* \in \mathbb{R}^{p\times T}$ has rank at most r, and the noise matrix $\mathbf{W}$ has i.i.d. entries that are zero-mean and σ-subGaussian. Then any solution to the program (3) with $\lambda_n = 10\sigma\sqrt{\gamma_{\max}(\widehat{\mathbf{\Sigma}})}\left(\sqrt{\frac{p+T}{n}}+\delta\right)$ satisfies the bound

$$\left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_2 \le 30\sqrt{2}\frac{\sigma\sqrt{\gamma_{\max}\left(\widehat{\mathbf{\Sigma}}\right)}}{\gamma_{\min}\left(\widehat{\mathbf{\Sigma}}\right)}\left(\sqrt{\frac{p+T}{n}}+\delta\right)$$

with probability at least $1-2e^{-2n\delta^2}$. Moreover, we have

$$\left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_{\mathrm{F}} \le 4\sqrt{2r}\left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_2 \text{ and } \left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_{\text{nuc}} \le 32r\left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_2.$$

- When $n > p$, the lower bound $\gamma_{\min}(\widehat{\mathbf{\Sigma}}) > 0$ cannot hold otherwise.
- However, even if the matrix $\mathbf{\Theta}^*$ were rank-one, it would have at least $p+T$ degrees of freedom, so this lower bound is unavoidable.

Proof

We first claim the curvature condition

$$\|\nabla \mathcal{L}_n(\boldsymbol{\Theta}^* + \boldsymbol{\Delta}) - \nabla \mathcal{L}_n(\boldsymbol{\Theta}^*)\|_2 \geqslant \gamma_{\min}(\widehat{\boldsymbol{\Sigma}})\|\boldsymbol{\Delta}\|_2.$$

We have $\nabla \mathcal{L}_n(\boldsymbol{\Theta}) = \frac{1}{n}\mathbf{Z}^{\mathrm{T}}(y - \mathbf{Z}\boldsymbol{\Theta})$, and hence $\nabla \mathcal{L}_n(\boldsymbol{\Theta}^* + \boldsymbol{\Delta}) - \nabla \mathcal{L}_n(\boldsymbol{\Theta}^*) = \widehat{\boldsymbol{\Sigma}}\boldsymbol{\Delta}$ where $\widehat{\boldsymbol{\Sigma}} = \frac{\mathbf{Z}^{\mathrm{T}}\mathbf{Z}}{n}$ is the sample covariance. Thus, it suffices to show that

$$\|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Delta}\|_2 \geq \gamma_{\min}(\widehat{\boldsymbol{\Sigma}})\|\boldsymbol{\Delta}\|_2 \quad \text{for all } \boldsymbol{\Delta} \in \mathbb{R}^{p\times T}.$$

For any vector $u \in \mathbb{R}^T$, we have $\|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Delta}u\|_2 \geq \gamma_{\min}(\widehat{\boldsymbol{\Sigma}})\|\boldsymbol{\Delta}u\|_2$, and thus

$$\|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Delta}\|_2 = \sup_{\|u\|_2=1} \|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Delta}u\|_2 \geq \gamma_{\min}(\widehat{\boldsymbol{\Sigma}}) \sup_{\|u\|_2=1} \|\boldsymbol{\Delta}u\|_2 = \gamma_{\min}(\widehat{\boldsymbol{\Sigma}})\|\boldsymbol{\Delta}\|_2,$$

which establishes the claim.

Proof

It remains to verify that the inequality $\|\nabla \mathcal{L}_n(\mathbf{\Theta}^*)\|_2 \leq \frac{\lambda_n}{2}$ holds with high probability under the stated choice of λ_n. For this model, we have $\nabla \mathcal{L}_n(\mathbf{\Theta}^*) = \frac{1}{n}\mathbf{Z}^{\mathrm{T}}\mathbf{W}$, where $\mathbf{W} \in \mathbb{R}^{n \times T}$ is a zero-mean matrix of i.i.d. σ-sub-Gaussian variates.

Proof

The next theorem is shown in High-Dimension Probability book.

Theorem 4.4.5 (Norm of matrices with sub-gaussian entries)

Let A be an $m \times n$ random matrix whose entries A_{ij} are independent mean-zero sub-gaussian random variables. Then, for any $t > 0$ we have

$$\|A\| \leq CK(\sqrt{m} + \sqrt{n} + t)$$

with probability at least $1 - 2\exp\left(-t^2\right)$. Here $K = \max_{i,j} \|A_{ij}\|_{\psi_2}$.

Definition: $\|X\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}\exp(X^2/t^2) \leqslant 2\}$.
We have

$$\mathbb{P}\left[\left\|\frac{1}{n}\mathbf{Z}^{\mathrm{T}}\mathbf{W}\right\|_2 \geq 5\sigma\sqrt{\gamma_{\max}(\widehat{\mathbf{\Sigma}})}\left(\sqrt{\frac{p+T}{n}} + \delta\right)\right] \leq 2e^{-2n\delta^2}$$

from which the validity of λ_n follows. Thus, the bound follows from Proposition 10.7.

Proof

Turning to the remaining bounds, with the given choice of λ_n, the cone inequality guarantees that $\left\|\widehat{\mathbf{\Delta}}_{\mathbb{M}^\perp}\right\|_{\textsf{nuc}} \leq 3\left\|\widehat{\mathbf{\Delta}}_{\overline{\mathbb{M}}}\right\|_{\textsf{nuc}}$.
Since any matrix in $\overline{\mathbb{M}}$ has rank at most $2r$, we conclude that

$$\|\widehat{\mathbf{\Delta}}\|_{\textsf{nuc}} \leq 4\|\widehat{\mathbf{\Delta}}_{\overline{\mathbb{M}}}\|_{\textsf{nuc}} \leq 4\sqrt{2r}\|\widehat{\mathbf{\Delta}}\|_{\mathrm{F}}.$$

We have

$$\|\widehat{\mathbf{\Delta}}\|_{\mathrm{F}}^2 = \langle\langle\widehat{\mathbf{\Delta}}, \widehat{\mathbf{\Delta}}\rangle\rangle \overset{(i)}{\leq} \|\widehat{\mathbf{\Delta}}\|_{\textsf{nuc}}\,\|\widehat{\mathbf{\Delta}}\|_2 \overset{(ii)}{\leq} 4\sqrt{2r}\|\widehat{\mathbf{\Delta}}\|_{\mathrm{F}}\|\widehat{\mathbf{\Delta}}\|_2,$$

where step (i) follows from Hölder's inequality, and step (ii) follows from our previous bound. Thereby completing the proof.

Outline

1. Matrix regression and applications
2. Analysis of nuclear norm regularization
3. Matrix compressed sensing
4. Bounds for phase retrieval
5. Multivariate regression with low-rank constraints
6. Matrix completion
7. Additive matrix decompositions

Netflix Challenge

					
Tom	4	?	5	?	?
Jerry	?	?	?	3	1
Alice	5	5	?	?	?
Bob	?	?	2	?	?
Qiwen	?	?	?	?	4

- Data: 480,189 users, 17,770 movies, 100,480,507 ratings (1-5).
- Year: 1998$\sim$ 2005.

Model: Matrix completion

Consider the matrix regression: observations are of the form

$$y_i = \langle\langle \mathbf{X}_i, \mathbf{\Theta}^* \rangle\rangle + w_i,$$

where $\mathbf{X}_i \in \mathbb{R}^{d_1 \times d_2}$ is a sparse mask matrix, zero everywhere except for a single randomly chosen entry $(a(i), b(i))$, where it is equal to $\sqrt{d_1 d_2}$.

- Let us now clarify why we chose to use rescaled mask matrices $\mathbf{X}_i$-that is, equal to $\sqrt{d_1 d_2}$ instead of 1 in their unique non-zero entry. With this choice, we have the convenient relation

$$\mathbb{E}\left[\frac{\|\mathfrak{X}_n(\mathbf{\Theta}^*)\|_2^2}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\langle\langle \mathbf{X}_i, \mathbf{\Theta}^* \rangle\rangle^2\right] = \|\mathbf{\Theta}^*\|_{\mathrm{F}}^2,$$

  using the fact that each entry of $\mathbf{\Theta}^*$ is picked out with probability $(d_1 d_2)^{-1}$.

Matrix incoherence condition

Consider the singular value decomposition $\mathbf{\Theta}^* = \mathbf{UDV}^{\mathrm{T}}$, where

$$\left\| \mathbf{U}\mathbf{U}^{\mathrm{T}} - \frac{r}{d_1}\mathbf{I}_{d_1 \times d_1} \right\|_{\max} \leq \mu \frac{\sqrt{r}}{d_1}, \quad \left\| \mathbf{V}\mathbf{V}^{\mathrm{T}} - \frac{r}{d_2}\mathbf{I}_{d_2 \times d_2} \right\|_{\max} \leq \mu \frac{\sqrt{r}}{d_2},$$

where $\mu > 0$ is the incoherence parameter.

- Each entry of matrix $\mathbf{U} \in \mathbb{R}^{d_1 \times r}$ would have magnitude of the order $1/\sqrt{d_1}$. As a consequence, in this ideal case, each r-dimensional row of $\mathbf{U}$ would have Euclidean norm exactly $\sqrt{r/d_1}$. Similarly, the rows of $\mathbf{V}$ would have Euclidean norm $\sqrt{r/d_2}$ in the ideal case.
- Non-robustness property

$$\mathbf{\Gamma}^* = (1-\delta)\mathbf{Z}^* + \delta \mathbf{\Theta}^{\mathrm{bad}} \qquad \text{for some } \delta \in (0, 1]$$

As long as $\delta > 0$, vector $z = [0, 1, 1, \cdots, 1]$, and the associated matrix $\mathbf{Z}^* := (z \otimes z)/d.$, $\mathbf{\Theta}^{bad} := e_1 \otimes e_1$, then the matrix $\mathbf{\Gamma}^*$ has $e_1 \in \mathbb{R}^d$ as one of its eigenvectors, and so violates the incoherence conditions.

Spikiness ratio

More precisely, for any non-zero matrix $\boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2}$, we define the spikiness ratio

$$\alpha_{\mathrm{sp}}(\boldsymbol{\Theta}) = \frac{\sqrt{d_1 d_2}\|\boldsymbol{\Theta}\|_{\max}}{\|\boldsymbol{\Theta}\|_{\mathrm{F}}},$$

where $\|\cdot\|_{\max}$ denotes the elementwise maximum absolute value. By definition of the Frobenius norm, we have

$$\|\boldsymbol{\Theta}\|_{\mathrm{F}}^2 = \sum_{j=1}^{d_1} \sum_{k=1}^{d_2} \Theta_{jk}^2 \leq d_1 d_2 \|\boldsymbol{\Theta}\|_{\max}^2,$$

so that the spikiness ratio is lower bounded by 1 . On the other hand, it can also be seen that $\alpha_{\mathrm{sp}}(\boldsymbol{\Theta}) \leq \sqrt{d_1 d_2}$. In particular, for any $\delta \in [0, 1]$, we have $\alpha_{\mathsf{sp}}\ (\boldsymbol{\Gamma}^*) = \frac{\sqrt{d_1 d_2}\max\{(1-\delta)/d,\delta\}}{\sqrt{(d_1-1)(d_2-1)(1-\delta)^2/d^2+\delta^2}}$.

Theorem 10.17

Let $\mathfrak{X}_n : \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}^n$ be the random matrix completion operator formed by n i.i.d. samples of rescaled mask matrices $\mathbf{X}_i$. Then there are universal positive constants (c_1, c_2) such that

$$\left| \frac{1}{n} \frac{\|\mathfrak{X}_n(\mathbf{\Theta})\|_2^2}{\|\mathbf{\Theta}\|_{\mathrm{F}}^2} - 1 \right| \leq c_1 \alpha_{\mathrm{sp}}(\mathbf{\Theta}) \frac{\|\mathbf{\Theta}\|_{\mathrm{nuc}}}{\|\mathbf{\Theta}\|_{\mathrm{F}}} \sqrt{\frac{d \log d}{n}} + c_2 \alpha_{\mathrm{sp}}^2(\mathbf{\Theta}) \left(\sqrt{\frac{d \log d}{n}} + \delta \right)$$

for all non-zero $\mathbf{\Theta} \in \mathbb{R}^{d_1 \times d_2}$, uniformly with probability at least $1 - 2e^{-\frac{1}{2} d \log d - n\delta}$.

- Theorem establishes a form of restricted strong convexity for the random operator that underlies matrix completion. Denote $d = d_1 + d_2$.

Noisy matrix completion

Consider the model

$$\widetilde{y}_i = \Theta_{a(i),b(i)} + \frac{w_i}{\sqrt{d_1 d_2}}. \tag{4}$$

Given n i.i.d. samples $\widetilde{y}_i$ from the noisy linear model (4), consider the nuclear norm regularized estimator

$$\widehat{\mathbf{\Theta}} \in \arg \min_{\|\mathbf{\Theta}\|_{\max} \le \frac{\alpha}{\sqrt{d_1 d_2}}} \left\{ \frac{1}{2n} \sum_{i=1}^{n} d_1 d_2 \left\{ y_i - \Theta_{a(i),b(i)} \right\}^2 + \lambda_n \|\mathbf{\Theta}\|_{\mathrm{nuc}} \right\} \tag{5}$$

where Theorem 10.17 motivates the addition of the extra side constraint on the infinity norm of $\mathbf{\Theta}$.

Error bound

Corollary 10.18

Consider the observation model (4) for a matrix $\mathbf{\Theta}^*$ with rank at most r, elementwise bounded as $\|\mathbf{\Theta}^*\|_{\max} \leq \alpha/\sqrt{d_1 d_2}$, and i.i.d. additive noise variables $\{w_i\}_{i=1}^n$ that satisfy the Bernstein condition with parameters (σ, b). Given a sample size $n > \frac{100b^2}{\sigma^2} d \log d$, if we solve the program (5) with $\lambda_n^2 = 25\frac{\sigma^2 d \log d}{n} + \delta^2$ for some $\delta \in \left(0, \frac{\sigma^2}{2b}\right)$, then any optimal solution $\widehat{\mathbf{\Theta}}$ satisfies the bound

$$\left\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}^*\right\|_{\mathrm{F}}^2 \leq c_1 \max\left\{\sigma^2, \alpha^2\right\} r \left\{\frac{d \log d}{n} + \delta^2\right\}.$$

with probability at least $1 - e^{-\frac{n\delta^2}{16d}} - 2e^{-\frac{1}{2} d \log d - n\delta}$.

Proof of Corollary 10.18

We first verify that the good event $\mathbb{G}(\lambda_n) = \left\{ \|\nabla \mathcal{L}_n(\mathbf{\Theta}^*)\|_2 \leq \frac{\lambda_n}{2} \right\}$ holds with high probability. The gradient of the least-squares objective is given by

$$\nabla \mathcal{L}_n(\mathbf{\Theta}^*) = \frac{1}{n} \sum_{i=1}^{n} (d_1 d_2) \frac{w_i}{\sqrt{d_1 d_2}} \mathbf{E}_i = \frac{1}{n} \sum_{i=1}^{n} w_i \mathbf{X}_i,$$

where we recall the rescaled mask matrices $\mathbf{X}_i := \sqrt{d_1 d_2} \mathbf{E}_i$. From our calculations in Example 6.18, we have

$$\mathbb{P}\left[\left\| \frac{1}{n} \sum_{i=1}^{n} w_i \mathbf{X}_i \right\|_2 \geq \epsilon \right] \leq 4d e^{-\frac{n\epsilon^2}{8d(\sigma^2 + b\epsilon)}} \leq 4d e^{-\frac{n\epsilon^2}{16 d \sigma^2}},$$

where the second inequality holds for any $\epsilon > 0$ such that $b\epsilon \leq \sigma^2$. Under the stated lower bound on the sample size, we are guaranteed that $b\lambda_n \leq \sigma^2$, from which it follows that the event $\mathbb{G}(\lambda_n)$ holds with the claimed probability.

Proof of Corollary 10.18

Next we use Theorem 10.17 to verify a variant of the restricted strong convexity condition.

Under the event $\mathbb{G}(\lambda_n)$, Proposition 9.13 implies that the error matrix $\widehat{\boldsymbol{\Delta}} = \widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}^*$ satisfies the constraint $\|\widehat{\boldsymbol{\Delta}}\|_{\text{nuc}} \leq 4\|\widehat{\boldsymbol{\Delta}}_{\bar{\mathbb{M}}}\|_{\text{nuc}}$. As noted earlier, any matrix in $\bar{\mathbb{M}}$ has rank at most $2r$, whence $\|\widehat{\boldsymbol{\Delta}}\|_{\text{nuc}} \leq 4\sqrt{2r}\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}$. By construction, we also have $\|\widehat{\boldsymbol{\Delta}}\|_{\max} \leq \frac{2\alpha}{\sqrt{d_1 d_2}}$.

Putting together the pieces, Theorem 10.17 implies that, with probability at least $1 - 2e^{-\frac{1}{2}d\log d - n\delta}$, the observation operator $\mathfrak{X}_n$ satisfies the lower bound

$$\begin{aligned}
\frac{\left\|\mathfrak{X}_n(\widehat{\boldsymbol{\Delta}})\right\|_2^2}{n} &\geq \|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2 - 8\sqrt{2}c_1\alpha\sqrt{\frac{rd\log d}{n}}\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}} - 4c_2\alpha^2\left(\sqrt{\frac{d\log d}{n}} + \delta\right)^2 \\
&\geq \|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}\left\{\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}} - 8\sqrt{2}c_1\alpha\sqrt{\frac{rd\log d}{n}}\right\} - 8c_2\alpha^2\left(\frac{d\log d}{n} + \delta^2\right)
\end{aligned}$$

Proof of Corollary 10.18

In order to complete the proof using this bound, we only need to consider two possible cases.

Case 1: On one hand, if either

$$\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}} \leq 16\sqrt{2}c_1\alpha\sqrt{\frac{rd\log d}{n}} \text{ or } \|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2 \leq 64c_2\alpha^2\left(\frac{d\log d}{n}+\delta^2\right),$$

then the claim follows.

Case 2: Otherwise, we must have

$$\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}} - 8\sqrt{2}c_1\alpha\sqrt{\frac{rd\log d}{n}} > \frac{\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}}{2} \text{ and } 8c_2\alpha^2\left(\frac{d\log d}{n}+\delta^2\right) < \frac{\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2}{4},$$

and hence the lower bound (10.44) implies that

$$\frac{\left\|\mathfrak{X}_n(\widehat{\boldsymbol{\Delta}})\right\|_2^2}{n} \geq \frac{1}{2}\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2 - \frac{1}{4}\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2 = \frac{1}{4}\|\widehat{\boldsymbol{\Delta}}\|_{\mathrm{F}}^2.$$

This is the required restricted strong convexity condition, let $\kappa = 1/2$ and by Proposition 10.6, so the proof is then complete.

Proof of Theorem 10.17

Given the invariance of the inequality to rescaling, we may assume without loss of generality that $\|\boldsymbol{\Theta}\|_{\mathrm{F}} = 1$. For given positive constants (α, ρ), define the set
$\mathbb{S}(\alpha, \rho) = \{\boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2} \mid \|\boldsymbol{\Theta}\|_{\mathrm{F}} = 1, \ \|\boldsymbol{\Theta}\|_{\max} \leq \frac{\alpha}{\sqrt{d_1 d_2}} \text{ and } \|\boldsymbol{\Theta}\|_{\mathrm{nuc}} \leq \rho\}$,
as well as the associated random variable
$Z(\alpha, \rho) := \sup_{\boldsymbol{\Theta} \in \mathcal{S}(\alpha, \rho)} \left| \frac{1}{n} \|\mathfrak{X}_n(\boldsymbol{\Theta})\|_2^2 - 1 \right|$. We begin by showing that there are universal constants (c_1, c_2) such that

$$\mathbb{P}\left[Z(\alpha, \rho) \geq \frac{c_1}{4} \alpha \rho \sqrt{\frac{d \log d}{n}} + \frac{c_2}{4} \left(\alpha \sqrt{\frac{d \log d}{n}} \right)^2 \right] \leq e^{-d \log d}.$$

Here our choice of the rescaling by $1/4$ is for later theoretical convenience.

Proof of Theorem 10.17

Concentration around mean: Note $F_{\boldsymbol{\Theta}}(\mathbf{X}) := \langle\langle \boldsymbol{\Theta}, \mathbf{X} \rangle\rangle^2$, we can write

$$Z(\alpha, \rho) = \sup_{\boldsymbol{\Theta} \in \mathbb{S}(\alpha,\rho)} \left| \frac{1}{n} \sum_{i=1}^{n} F_{\boldsymbol{\Theta}}(\mathbf{X}_i) - \mathbb{E}\left[F_{\boldsymbol{\Theta}}(\mathbf{X}_i)\right] \right|,$$

We need to bound $\|F_{\boldsymbol{\Theta}}\|_{\max}$ and $\operatorname{var}(F_{\boldsymbol{\Theta}}(\mathbf{X}))$ uniformly over the class. For any rescaled mask matrix $\mathbf{X}$ and parameter matrix $\boldsymbol{\Theta} \in \mathbb{S}(\alpha, \rho)$, we have

$$|F_{\boldsymbol{\Theta}}(\mathbf{X})| \leq \|\boldsymbol{\Theta}\|_{\max}^2 \|\mathbf{X}\|_1^2 \leq \frac{\alpha^2}{d_1 d_2} d_1 d_2 = \alpha^2,$$

Turning to the variance, we have

$$\operatorname{var}(F_{\boldsymbol{\Theta}}(\mathbf{X})) \leq \mathbb{E}\left[F_{\boldsymbol{\Theta}}^2(\mathbf{X})\right] \leq \alpha^2 \mathbb{E}\left[F_{\boldsymbol{\Theta}}(\mathbf{X})\right] = \alpha^2,$$

a bound which holds for any $\boldsymbol{\Theta} \in \mathbb{S}(\alpha, \rho)$. Consequently, applying the bound (3.86) with $\epsilon = 1$ and $t = d \log d / n$, we conclude that there are universal constants (c_1, c_2) such that

$$\mathbb{P}\left[Z(\alpha, \rho) \geq 2\mathbb{E}[Z(\alpha, \rho)] + \frac{c_1}{8} \alpha \sqrt{\frac{d \log d}{n}} + \frac{c_2}{4} \alpha^2 \frac{d \log d}{n}\right] \leq e^{-d \log d}.$$

Proof of Theorem 10.17

Bounding the expectation: It remains to bound the expectation. By Rademacher symmetrization (see Proposition 4.11), we have

$$\begin{aligned}\mathbb{E}[Z(\alpha,\rho)] &\leq 2\mathbb{E}\left[\sup_{\boldsymbol{\Theta}\in\mathbb{S}(\alpha,\rho)}\left|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\left\langle\!\left\langle\mathbf{X}_i,\boldsymbol{\Theta}\right\rangle\!\right\rangle^2\right|\right] \\ &\overset{\text{(ii)}}{\leq} 4\alpha\mathbb{E}\left[\sup_{\boldsymbol{\Theta}\in\mathcal{S}(\alpha,\rho)}\left|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\left\langle\!\left\langle\mathbf{X}_i,\boldsymbol{\Theta}\right\rangle\!\right\rangle\right|\right],\end{aligned} \tag{6}$$

where inequality (ii) follows from the Ledoux-Talagrand contraction inequality (5.61) for Rademacher processes, using the fact that $|\langle\!\langle\boldsymbol{\Theta},\mathbf{X}_i\rangle\!\rangle|\leq\alpha$ for all pairs $(\boldsymbol{\Theta},\mathbf{X}_i)$.
Next we apply Hölder's inequality to bound the remaining term: since $\|\boldsymbol{\Theta}\|_{\text{nuc}}\leq\rho$ for any $\boldsymbol{\Theta}\in\mathbb{S}(\alpha,\rho)$, we have

$$\mathbb{E}\left[\sup_{\mathbb{S}(\alpha,\rho)}\left|\left\langle\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\mathbf{X}_i,\boldsymbol{\Theta}\right\rangle\right|\right]\leq\rho\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\mathbf{X}_i\right\|_2\right].$$

Proof of Theorem 10.17

Finally, note that each matrix $\varepsilon_i \mathbf{X}_i$ is zero-mean, has its operator norm upper bounded as $\|\varepsilon_i \mathbf{X}_i\|_2 \leq \sqrt{d_1 d_2} \leq d$, and its variance bounded as

$$\|\mathrm{var}\,(\varepsilon_i \mathbf{X}_i)\|_2 = \frac{1}{d_1 d_2} \|d_1 d_2 (1 \otimes 1)\|_2 = \sqrt{d_1 d_2}.$$

Consequently, the result of Exercise 6.10 implies that

$$\mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{X}_i\right\|_2 \geq \delta\right] \leq 2d \exp\left\{-\frac{n\delta^2}{2d(1+\delta)}\right\}.$$

Next, applying Exercise 2.8(a) with $C = 2d, v^2 = \frac{d}{n}$ and $B = \frac{d}{n}$, we find that

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{X}_i\right\|_2\right] \leq 2\sqrt{\frac{d}{n}}(\sqrt{\log(2d)} + \sqrt{\pi}) + \frac{4d\log(2d)}{n} \overset{\text{(i)}}{\leq} 16\sqrt{\frac{d\log d}{n}}.$$

Here the inequality (i) uses the fact that $n > d \log d$. We conclude that

$$\mathbb{E}[Z(\alpha,\rho)] \leq \frac{c_1}{16}\alpha\rho\sqrt{\frac{d\log d}{n}}.$$

Proof of Theorem 10.17

Extension via peeling: Let $\mathbb{B}_F(1)$ denote the Frobenius ball of norm one in $\mathbb{R}^{d_1 \times d_2}$, and let $\mathcal{E}$ be the event that the bound (10.41) is violated for some $\mathbf{\Theta} \in \mathbb{B}_F(1)$. For $k, \ell = 1, 2, \ldots$, let us define the sets

$$\mathbb{S}_{k,\ell} := \left\{ \mathbf{\Theta} \in \mathbb{B}_F(1) \mid 2^{k-1} \le d \|\mathbf{\Theta}\|_{\max} \le 2^k \text{ and } 2^{\ell-1} \le \|\mathbf{\Theta}\|_{\mathrm{nuc}} \le 2^{\ell} \right\},$$

and let $\mathcal{E}_{k,\ell}$ be the event that the bound (10.41) is violated for some $\mathbf{\Theta} \in \mathbb{S}_{k,\ell}$. We first claim that

$$\mathcal{E} \subseteq \bigcup_{k,\ell=1}^{M} \mathcal{E}_{k,\ell}, \quad \text{where } M = \lceil \log d \rceil. \tag{7}$$

Indeed, for any matrix $\mathbf{\Theta} \in \mathbb{S}(\alpha, \rho)$, we have

$$\|\mathbf{\Theta}\|_{\mathrm{nuc}} \ge \|\mathbf{\Theta}\|_{\mathrm{F}} = 1 \quad \text{and} \quad \|\mathbf{\Theta}\|_{\mathrm{nuc}} \le \sqrt{d_1 d_2} \|\mathbf{\Theta}\|_{\mathrm{F}} \le d.$$

Thus, we may assume that $\|\mathbf{\Theta}\|_{\mathrm{nuc}} \in [1, d]$ without loss of generality. For any matrix of Frobenius norm one, we have $d\|\mathbf{\Theta}\|_{\max} \ge \sqrt{d_1 d_2} \|\mathbf{\Theta}\|_{\max} \ge 1$ and $d\|\mathbf{\Theta}\|_{\max} \le d$, showing that we may also assume that $d\|\mathbf{\Theta}\|_{\max} \in [1, d]$. Then (7) holds for $k, \ell = 1, 2 \ldots, M$, with $M = \lceil \log d \rceil$.

Proof of Theorem 10.17

Next, for $\alpha = 2^k$ and $\rho = 2^\ell$, define the event

$$\widetilde{\mathcal{E}}_{k,\ell} := \left\{ Z(\alpha, \rho) \geq \frac{c_1}{4}\alpha\rho\sqrt{\frac{d \log d}{n}} + \frac{c_2}{4}\left(\alpha\sqrt{\frac{d \log d}{n}}\right)^2 \right\}.$$

We claim that $\mathcal{E}_{k,\ell} \subseteq \widetilde{\mathcal{E}}_{k,\ell}$. Indeed, if event $\mathcal{E}_{k,\ell}$ occurs, then there must exist some $\mathbf{\Theta} \in \mathbb{S}_{k,\ell}$ such that

$$\begin{aligned}
\left| \frac{1}{n}\|\mathfrak{X}_n(\mathbf{\Theta})\|_2^2 - 1 \right| &\geq c_1 d\|\mathbf{\Theta}\|_{\max}\|\mathbf{\Theta}\|_{\mathrm{nuc}}\sqrt{\frac{d \log d}{n}} + c_2\left(d\|\mathbf{\Theta}\|_{\max}\sqrt{\frac{d \log d}{n}}\right) \\
&\geq c_1 2^{k-1}2^{\ell-1}\sqrt{\frac{d \log d}{n}} + c_2\left(2^{k-1}\sqrt{\frac{d \log d}{n}}\right)^2 \\
&\geq \frac{c_1}{4}2^k 2^\ell\sqrt{\frac{d \log d}{n}} + \frac{c_2}{4}\left(2^k\sqrt{\frac{d \log d}{n}}\right)^2
\end{aligned}$$

showing that $\widetilde{\mathcal{E}}_{k,\ell}$ occurs.

Proof of Theorem 10.17

Putting together the pieces, we have

$$\mathbb{P}[\mathcal{E}] \stackrel{(\text{i})}{\leq} \sum_{k,\ell=1}^{M} \mathbb{P}\left[\widetilde{\mathcal{E}}_{k,\ell}\right] \stackrel{(\text{ii})}{\leq} M^2 e^{-d \log d} \stackrel{(\text{iii})}{\leq} e^{-\frac{1}{2} d \log d},$$

where inequality (i) follows from the union bound applied to the inclusion $\mathcal{E} \subseteq \bigcup_{k,\ell=1}^{M} \widetilde{\mathcal{E}}_{k,\ell}$; inequality (ii) is a consequence of the earlier tail bound (10.46); and inequality (iii) follows since $\log M^2 = 2 \log \log d \leq \frac{1}{2} d \log d$.

Outline

Additive matrix decomposition

Consider a pair of matrices : low-rank matrix $\mathbf{\Lambda}^*$ and sparse matrix $\mathbf{\Gamma}^*$, and suppose that we observe a vector $y \in \mathbb{R}^n$ of the form

$$y = \mathfrak{X}_n \left(\mathbf{\Lambda}^* + \mathbf{\Gamma}^* \right) + w,$$

where $\mathfrak{X}_n$ is a known linear observation operator, mapping matrices in $\mathbb{R}^{d_1 \times d_2}$ to a vector in $\mathbb{R}^n$. Consider the following estimator:

$$(\widehat{\mathbf{\Gamma}}, \widehat{\mathbf{\Lambda}}) = \arg \min_{\substack{\mathbf{\Gamma} \in \mathbb{R}^{d_1 \times d_2} \\ \|\mathbf{\Lambda}\|_{\max} \leq \frac{\alpha}{\sqrt{d_1 d_2}}}} \left\{ \frac{1}{2} \|\mathbf{Y} - (\mathbf{\Gamma} + \mathbf{\Lambda})\|_{\mathrm{F}}^2 + \lambda_n \left(\|\mathbf{\Gamma}\|_1 + \omega_n \|\mathbf{\Lambda}\|_{\mathrm{nuc}} \right) \right\}.$$

It is parameterized by two regularization parameters, namely λ_n and ω_n. Define the squared Frobenius norm error

$$e^2 \left(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}^*, \widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}^* \right) := \left\| \widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}^* \right\|_{\mathrm{F}}^2 + \left\| \widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}^* \right\|_{\mathrm{F}}^2.$$

Corollary 10.22

Suppose that we solve the convex program with parameters

$$\lambda_n \geq 2\|\mathbf{W}\|_{\max} + 4\frac{\alpha}{\sqrt{d_1 d_2}} \text{ and } \omega_n \geq \frac{2\|\mathbf{W}\|_2}{\lambda_n}.$$

Then there are universal constants c_j such that for any matrix pair $(\mathbf{\Lambda}^*, \mathbf{\Gamma}^*)$ with $\|\mathbf{\Lambda}^*\|_{\max} \leq \frac{\alpha}{\sqrt{d_1 d_2}}$ and for all integers $r = 1, 2, \ldots, \min\{d_1, d_2\}$ and $s = 1, 2, \ldots, (d_1 d_2)$, the squared Frobenius error is upper bounded as

$$c_1 \omega_n^2 \lambda_n^2 \left\{ r + \frac{1}{\omega_n \lambda_n} \sum_{j=r+1}^{\min\{d_1, d_2\}} \sigma_j\left(\mathbf{\Lambda}^*\right) \right\} + c_2 \lambda_n^2 \left\{ s + \frac{1}{\lambda_n} \sum_{(j,k) \notin S} \left|\mathbf{\Gamma}_{jk}^*\right| \right\},$$

where S is an arbitrary subset of matrix indices of cardinality at most s.

Proof outline

Applying Theorem 9.19. Doing so requires three steps:

- Verifying a form of restricted strong convexity;
- Verifying the validity of the regularization parameters;
- Computing the subspace Lipschitz constant from Definition 9.18.

Proof of Corollary 10.22

We begin with restricted strong convexity. Define the two matrices $\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}} = \widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}^*$ and $\mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}} := \widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}^*$, By expanding out the quadratic form, we find that the first-order error in the Taylor series is given by

$$\mathcal{E}_n\left(\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}, \mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right) = \frac{1}{2}\left\|\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}} + \mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right\|_{\mathrm{F}}^2 = \frac{1}{2}\underbrace{\left\{\left\|\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}\right\|_{\mathrm{F}}^2 + \left\|\mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right\|_{\mathrm{F}}^2\right\}}_{e^2\left(\mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}, \mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}\right)} + \left\langle\left\langle \mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}, \mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}} \right\rangle\right\rangle.$$

By the triangle inequality and the construction of our estimator, we have

$$\left\|\mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right\|_{\mathsf{max}} \le \|\widehat{\mathbf{\Lambda}}\|_{\mathsf{max}} + \|\mathbf{\Lambda}^*\|_{\mathsf{max}} \le \frac{2\alpha}{\sqrt{d_1 d_2}}.$$

Combined with Hölder's inequality, we see that

$$\mathcal{E}_n\left(\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}, \mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right) \ge \frac{1}{2} e^2\left(\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}, \mathbf{\Delta}_{\widehat{\mathbf{\Lambda}}}\right) - \frac{2\alpha}{\sqrt{d_1 d_2}}\left\|\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}\right\|_1,$$

so that restricted strong convexity holds with $\kappa = 1$, since the remaining term proportional to $\left\|\mathbf{\Delta}_{\widehat{\mathbf{\Gamma}}}\right\|_1$, the proof of Theorem 9.19 shows that it can be absorbed without any consequence as long as $\lambda_n \ge \frac{4\alpha}{\sqrt{d_1 d_2}}$.

Proof of Corollary 10.22

Verifying event $\mathbb{G}(\lambda_n)$: A straightforward calculation gives $\nabla \mathcal{L}_n(\mathbf{\Gamma}^*, \mathbf{\Lambda}^*) = (\mathbf{W}, \mathbf{W})$. From the dual norm pairs given in Table 9.1, we have

$$\Phi^*_{\omega_n}(\nabla \mathcal{L}_n(\mathbf{\Gamma}^*, \mathbf{\Lambda}^*)) = \max\left\{ \|\mathbf{W}\|_{\max}, \frac{\|\mathbf{W}\|_2}{\omega_n} \right\},$$

so that the choices guarantee that $\lambda_n \geq 2\Phi^*_{\omega_n}(\nabla \mathcal{L}_n(\mathbf{\Gamma}^*, \mathbf{\Lambda}^*))$.

Proof of Corollary 10.22

Choice of model subspaces: For any subset S of matrix indices of cardinality at most s, define the subset $\mathbb{M}(S) := \left\{\mathbf{\Gamma} \in \mathbb{R}^{d_1 \times d_2} \mid \Gamma_{ij} = 0 \text{ for all } (i,j) \notin S\right\}$. Similarly, for any $r = 1, \ldots, \min\{d_1, d_2\}$, let $\mathbb{U}_r$ and $\mathbb{V}_r$ be (respectively) the subspaces spanned by the top r left and right singular vectors of $\mathbf{\Lambda}^*$, and recall the subspaces $\bar{\mathbb{M}}(\mathbb{U}_r, \mathbb{V}_r)$ and $\mathbb{M}^\perp(\mathbb{U}_r, \mathbb{V}_r)$. We are then guaranteed that the regularizer $\Phi_{\omega_n}(\mathbf{\Gamma}, \mathbf{\Lambda}) = \|\mathbf{\Gamma}\|_1 + \omega_n \|\mathbf{\Lambda}\|_{\text{nuc}}$ is decomposable with respect to the model subspace
$\mathbb{M} := \mathbb{M}(S) \times \bar{\mathbb{M}}(\mathbb{U}_r, \mathbb{V}_r)$ and deviation space $\mathbb{M}^\perp(S) \times \mathbb{M}^\perp(\mathbb{U}_r, \mathbb{V}_r)$. It then remains to bound the subspace Lipschitz constant. We have

$$\Psi(\mathbb{M}) = \sup_{(\mathbf{\Gamma},\Lambda) \in \mathbb{M}(S) \times \mathbb{M}(\mathbb{U}_r, \mathbb{V}_r)} \frac{\|\mathbf{\Gamma}\|_1 + \omega_n \|\mathbf{\Lambda}\|_{nuc}}{\sqrt{\|\mathbf{\Gamma}\|_{\mathrm{F}}^2 + \|\mathbf{\Lambda}\|_{\mathrm{F}}^2}} \leq \sup_{(\mathbf{\Gamma}, \mathbf{\Lambda})} \frac{\sqrt{s}\|\mathbf{\Gamma}\|_{\mathrm{F}} + \omega_n \sqrt{2r} \|\mathbf{\Lambda}\|_{\mathrm{F}}}{\sqrt{\|\mathbf{\Gamma}\|_{\mathrm{F}}^2 + \|\mathbf{\Lambda}\|_{\mathrm{F}}^2}}$$
$$\leq \sqrt{s} + \omega_n \sqrt{2r}.$$

Putting together the pieces, the overall claim now follows as a corollary of Theorem 9.19.