Robust Linear Algebra

Dimitris Bertsimas, Thodoris Koukouvinos

[a] *Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, 02139, MA, USA*
[b] *Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, 02139, MA, USA*

Abstract

We propose a robust optimization (RO) framework that immunizes some of the central linear algebra problems in the presence of data uncertainty. Namely, we formulate linear systems, matrix inversion, eigenvalues-eigenvectors and matrix factorization under uncertainty, as robust optimization problems using appropriate descriptions of uncertainty. The resulting optimization problems are computationally tractable and scalable. We show in theory that RO improves the relative error of the linear system by reducing the condition number of the underlying matrix. Moreover, we provide empirical evidence showing that the proposed approach outperforms state of the art methods for linear systems and matrix inversion, when applied on ill-conditioned matrices. We show that computing eigenvalues-eigenvectors under RO, corresponds to solving linear systems that are better conditioned than the nominal and illustrate with numerical experiments that the proposed approach is more accurate than the nominal, when perturbing ill-conditioned matrices. Finally, we demonstrate empirically the benefit of the robust Cholesky factorization over the nominal.

Keywords: Robustness and sensitivity analysis, Linear systems, Matrix inversion, Eigenvalues, Choleksy

1. Introduction

Current methods for solving linear algebra problems are sensitive to data perturbations. These methods have been developed assuming that the input data is exact. In many applications though, this is not the case and often there are accuracy issues especially for the case of nearly singular matrices. Motivated by these facts, we propose a framework capable of solving some

of the central linear algebra problems, that improves accuracy under data uncertainty. Although applicable in all cases, the regime where the proposed approach is most advantageous is that of ill-conditioned problems.
Our main tool in building this framework is RO (see (Bertsimas and den Hertog, 2022), (Ben-Tal et al., 2009) for details). Under RO, the way we model uncertainty is the notion of an uncertainty set $\mathcal{U}$. A linear RO problem is then defined as

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \quad \forall \boldsymbol{A} \in \mathcal{U}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned}$$

where $\mathcal{U}$ is a convex set. Let $\|\boldsymbol{A}\|_F = \sqrt{\sum_{i,j} A_{ij}^2}$ be the Frobenius norm. As an example, $\mathcal{U} = \{\boldsymbol{A} : \|\boldsymbol{A} - \overline{\boldsymbol{A}}\|_F \leq \rho\}$, where $\overline{\boldsymbol{A}}$ is the nominal matrix. In this paper, we formulate and solve some of the central linear algebra problems under uncertainty, using RO. Then, for norm bounded uncertainty sets we derive the robust counterpart (RC), which is a tractable optimization problem, see (Bertsimas and den Hertog, 2022).

Literature Review

Linear Systems

A lot of work is focusing on perturbation analysis in linear systems, where the goal there is to solve the linear system for a nominal matrix and provide bounds on how much the solution would change if some entries of the input matrix change. More details about it can be found in (Datta, 2010), (Trefethen and Bau III, 1997), (Higham, 2002), (Björck, 1991), (Eldén, 1980), (Fierro and Bunch, 1994). A lot of work has been done on the uncertain least squares problem, which has the same formulation as the linear system. One common approach in the literature is solving the problem through some matrix factorization. (Golub and Van Loan, 1980) addressed the least squares problem under uncertainty in both the input matrix and right-hand side vector, leveraging the Singular Value Decomposition (SVD) and also the Generalized Singular Value Decomposition. (Doyle et al., 1994) performed a robust analysis for linear systems utilizing again the SVD. Finally, we note that a special case of robust linear systems that has been studied a lot is interval linear systems, where the entries of the input matrix are unknown

and assumed to lie on independent interval uncertainty sets (Fiedler et al., 2006), (Kreinovich et al., 2013). We further discuss this case in Section 2.2.

(El Ghaoui and Lebret, 1997) utilized optimization problems, to find robust solutions for total least-squares. More recently, Xu et al. (2010) and (Bertsimas and Copenhaver, 2018) formulated the robust least squares problem as a computationally tractable optimization problem, utilizing RO. Further, (Zhen and den Hertog, 2017) addressed linear systems under uncertainty, while focusing mostly on the characterization of the robust solution set.

Widely used methods for linear systems include Gaussian elimination and iterative methods. Two state of the art iterative methods are GMRES (Saad and Schultz, 1986) and MINRES (Paige and Saunders, 1975), where the objective in both is to minimize the norm error, leveraging sophisticated methodologies, for example Krylov subspaces. The Conjugate Gradient method (Daniel, 1967) has also been used for linear systems. The method of choice is Gaussian elimination, since it can solve the linear system exactly as long as the input matrix is not singular. Further, (Fong and Saunders, 2011) and (Choi, 2006) developed iterative methods for singular systems.

Matrix Inversion

Most of the linear algebra community has focused on finding the inverse of the nominal matrix and calculating bounds on how much it would change if the entries of the nominal matrix are perturbed. Usually, the bounds come from the condition number of the matrix, where for large condition numbers significant changes are expected. (El Ghaoui, 2002) provided a bound on the change of the matrix inverse through semidefinite optimization, in the case of structured perturbations. (Hager, 1989) proposed calculating the inverse of a matrix under uncertainty utilizing the Sherman-Morrison-Woodbury formula. The main drawback of that approach is that a large amount of computational time is needed, since in order to calculate the matrix inverse, some other inverses need to be calculated first.

Eigenvalues-Eigenvectors

The problem of computing eigenvalues and eigenvectors (eigen-pairs) arises often in many applications including Principal Component Analysis and Spectral Clustering among many others. The eigenvalue problem for a square symmetric matrix of dimension n, requires the solution of a high degree

polynomial, which for higher values of n becomes intractable in practice. Iterative algorithms, such as the power method (Datta, 2010), have been proposed and are capable of finding the largest or smallest eigenvalue of a matrix. Currently, the LOBPCG method (Knyazev, 2001) is used in practice, which is based on an iterative scheme leveraging the conjugate gradient algorithm along with a three-term recurrence. The latter includes at each iteration the approximation of the current eigenvector, the current residual and the difference between the current and previous eigenvector approximations. (Scherzinger and Dohrmann, 2008) addressed the robust eigenvalue problem for 3×3 symmetric matrices. Finally, (Xu et al., 2010) and (Netrapalli et al., 2014) addressed the robust PCA problem, where the main goal is to find the largest robust eigen-pairs.

Matrix Factorizations

Matrix factorization is often used in linear algebra, including Cholesky, QR and SVD. Some work has been done on robust nonnegative matrix factorization (Huang et al., 2014) as well as on robust PCA (Candès et al., 2011). On the other hand, not much progress has been made on robustyfying the common linear algebra factorizations. (Ogita and Oishi, 2012) addressed robust Cholesky, though not utilizing RO.

Contributions

Our main contributions can be summarized as follows:

1. We address linear systems under uncertainty, using the formulations developed in (Zhen and den Hertog, 2017), (Bertsimas and Copenhaver, 2018). We illustrate that the robust solution is also an exact solution of multiple nearby linear systems. Moreover, we show that under RO the condition number of the underlying matrix for the linear system is reduced over that of the nominal. In the numerical experiments, we show that the robust approach outperforms state of the art methods for linear systems on ill-conditioned matrices.
2. We apply the framework introduced for robust linear systems, in the problem of robust matrix inversion and illustrate the advantage of this method when inverting nearly singular matrices, compared to state of the art approaches.

3. We propose an alternating optimization approach for computing robust eigen-pairs. We show that each robust eigen-pair is an exact eigen-pair on nearby matrices. Further, we show that the proposed method improves the relative error over the nominal eigenvectors by reducing the condition number of the underlying matrix. Finally, we demonstrate empirically that the robust eigen-pairs are more accurate than the nominal, on perturbations of nearly singular matrices.
4. We provide a framework for computing robust matrix factorizations, based on which we formulate the Cholesky factorization under uncertainty as a non-convex optimization problem. While not provably optimal, we demonstrate that high quality solutions obtained from the latter outperform the nominal Cholesky factorization in terms of accuracy on perturbed matrices.

The structure of the paper is as follows: In Section 2, we address linear systems under uncertainty. In Section 3, we address matrix inversion under uncertainty. In Section 4, we address eigenvalues under uncertainty, in Section 5 matrix factorization under uncertainty and finally in Section 6 we conclude.

We use bold faced characters such as $\boldsymbol{x}$ to represent vectors and bold faced capital letters such as $\boldsymbol{X}$ to represent matrices. We define $[n] = \{1, \ldots, n\}$. With $\boldsymbol{e}$ we denote the vector of ones and with $\boldsymbol{e}_i$ we denote the i-th unit vector. The norm $\|\boldsymbol{x}\|_p$ refers to the Euclidean norm for vectors, $\|\boldsymbol{x}\|_p = (\sum_i |x_i|^p)^{1/p}$. The norm $\|\boldsymbol{X}\|_F$ refers to the Frobenius norm for matrices, $\|\boldsymbol{X}\|_F = \sqrt{\sum_{i,j} X_{ij}^2}$. The norm $\|\boldsymbol{X}\|_1$ refers to the sum of the entries of the matrix in absolute value. The norm $\|\boldsymbol{X}\|_{\sigma_p}$ is defined as $\|\mu(\boldsymbol{X})\|_p$, where $\mu(\boldsymbol{X})$ denotes the vector of singular values of $\boldsymbol{X}$. The notation $|\boldsymbol{x}|$ refers to a vector containing the absolute value of each entry of $\boldsymbol{x}$. The inner product of matrices $\boldsymbol{A}, \boldsymbol{B}$ is defined as $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \text{tr}(\boldsymbol{A}^T \boldsymbol{B})$, where $\text{tr}(\boldsymbol{A})$ denotes the trace of matrix $\boldsymbol{A}$. The set $\mathcal{S}^n$ refers to all symmetric matrices of dimension n and $\mathcal{S}^n_+$ to those that are positive semi-definite. We use the symbols $\mathcal{U}_2, \mathcal{U}_1, \mathcal{U}_{11}, \mathcal{U}_\infty, \mathcal{U}_{\sigma_p}, \mathcal{V}$ to denote uncertainty sets. Finally, we use the notation $\mathcal{O}(g(x))$ to denote that a function $f(x)$ is of the order of $g(x)$, if there exists a constant C independent of x as well as some x_0, such that $f(x) \leq Cg(x)$, for all $x \geq x_0$.

All numerical experiments throughout the paper are performed on one Intel i9 2.3GHz CPU core with 16 GB RAM, using Gurobi version 9.0.2, implemented in Julia version 1.5.3 and Python version 3.7.7. The code uti-

lized for the numerical experiments can be found in the following github repository: `https://github.com/ThKoukouv/Robust-Linear-Algebra`.

2. Robust Linear Systems

In this section, we formulate the robust linear system (RLS) as a computationally tractable optimization problem, show some properties of the robust solution and also demonstrate the benefit over the nominal approach in the numerical experiments.

2.1. Formulation

For a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^n$ the linear system consists of finding a vector $\boldsymbol{x} \in \mathbb{R}^n$, such that $\boldsymbol{Ax} = \boldsymbol{b}$. In order to solve a classical linear system (CLS) one has to do pivots or in general a procedure known as Gaussian Elimination where one tries to create an upper triangular matrix from the input matrix and then solves the system with backward substitution. This procedure is the most widely used algorithm to solve a linear system and in the case of matrices that are non singular it is guaranteed to find the exact solution with a total of $\mathcal{O}(n^3)$ arithmetic operations.
The linear system can also be viewed as minimizing the norm error

$$\min_{\boldsymbol{x}} \ \|\boldsymbol{Ax} - \boldsymbol{b}\|_2. \tag{1}$$

When the entries of matrix $\boldsymbol{A}$ are subject to uncertainty, we need to be able to account for that uncertainty. We assume that we have additive perturbation for the matrix $\boldsymbol{A}$, that is $\boldsymbol{A} + \boldsymbol{\Delta A}$, and that the matrix $\boldsymbol{\Delta A}$ lies in an uncertainty set $\mathcal{U}$. The uncertainty set $\mathcal{U}$ should be able to capture uncertainty in the data and also result in a computationally tractable optimization problem. We utilize norm ball uncertainty sets, since they satisfy both properties. The first set that we consider is

$$\mathcal{U}_2 = \{\boldsymbol{\Delta A} \in \mathbb{R}^{n\times n} : \|\boldsymbol{\Delta A}\|_F \leq \rho\},$$

where ρ is an input parameter. In this case, each entry ΔA_{ij} is bounded by ρ, allowing each A_{ij} to change at most by ρ. Moreover, not all entries of the nominal matrix are allowed to take the worst case values. More precisely, if some entries of the nominal matrix are more uncertain, in which case ΔA_{ij} is larger, then others should be less uncertain so that the constraint $\|\boldsymbol{\Delta A}\|_F \leq \rho$ is satisfied. We can also describe uncertainty column-wise by

requiring that the norm of each column of $\boldsymbol{\Delta A}$ should be at most ρ. In this case the uncertainty set is

$$\mathcal{U}_1 = \{\boldsymbol{\Delta A} \in \mathbb{R}^{n\times n} : \|\boldsymbol{\Delta A}_i\|_2 \leq \rho,\ \forall i \in [n]\},$$

where $\boldsymbol{\Delta A}_i$ denotes the i-th column of $\boldsymbol{\Delta A}$. In this case, we allow for uncertainty in the entries of $\boldsymbol{A}$ but control the total amount of uncertainty within each column with the parameter ρ. Unlike the uncertainty set $\mathcal{U}_2$, here we can have multiple entries of ΔA_{ij} taking large values, as long as they are in different columns. For example, we can have $\boldsymbol{\Delta A}_i = \rho \boldsymbol{e}_i$. Observe that this choice of $\boldsymbol{\Delta A}$ does not belong in $\mathcal{U}_2$, since $\|\boldsymbol{\Delta A}\|_F = \sqrt{n}\rho$. Therefore, $\mathcal{U}_1$ is a more conservative uncertainty set than $\mathcal{U}_2$. We refer to Bertsimas and den Hertog (2022) for more discussion about the uncertainty sets. Finally, we consider an uncertainty set similar to $\mathcal{U}_1$ but with the ℓ_1 norm instead, that is

$$\mathcal{U}_{11} = \{\boldsymbol{\Delta A} \in \mathbb{R}^{n\times n} : \|\boldsymbol{\Delta A}_i\|_1 \leq \rho,\ \forall i \in [n]\}.$$

We follow a RO approach, which suggests solving Problem (1) over the maximum possible realization of uncertainty. The formulation is as follows:

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{\Delta A}\in\mathcal{U}} \|(\boldsymbol{A} + \boldsymbol{\Delta A})\boldsymbol{x} - \boldsymbol{b}\|_2, \tag{2}$$

where $\mathcal{U} \in \{\mathcal{U}_2, \mathcal{U}_1, \mathcal{U}_{11}\}$. Problem (2) can be viewed as an outer minimization problem over $\boldsymbol{x}$ and an inner maximization problem over $\boldsymbol{\Delta A}$. In fact, for the uncertainty sets that we consider, the inner problem has a closed form solution that we can obtain and then have an unconstrained minimization problem over $\boldsymbol{x}$. As has been shown in (El Ghaoui and Lebret, 1997), (Bertsimas and Copenhaver, 2018), in the context of robust regression, RLS with the proposed uncertainty sets can be reformulated as computationally tractable optimization problems.

Theorem 1 ((Bertsimas and Copenhaver, 2018),Xu et al. (2010)).

The RC of Problem (2) *with* $\mathcal{U} = \mathcal{U}_2$ *is the following problem:*

$$\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2 + \rho\|\boldsymbol{x}\|_2. \tag{3}$$

The RC of Problem (2) *with* $\mathcal{U} = \mathcal{U}_1$ *is the following problem:*

$$\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2 + \rho\|\boldsymbol{x}\|_1. \tag{4}$$

The RC of Problem (2) *with the* ℓ_1 *norm in the objective and* $\mathcal{U} = \mathcal{U}_{11}$ *is the following problem:*

$$\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_1 + \rho\|\boldsymbol{x}\|_1. \tag{5}$$

The above are the main approaches we utilize for RLS, but of course there can be further extensions in case of more complex uncertainty sets. The main benefit of the aforementioned uncertainty sets is that the resulting RCs are computationally tractable optimization problems. Specifically, Problem (5) is a LO and Problem (4) is a CQO, which can be solved efficiently with commercial solvers, i.e. Gurobi (Gurobi Optimization, LLC, 2021). We note that Bertsimas and Ma (2023) illustrated that Problem (3) with each norm term in the objective squared, also known as the Tikhonov regularization problem, is equivalent to the following robust problem

$$\min_{\boldsymbol{x}} \int_{\boldsymbol{\Delta A}\in\mathcal{U}} \|(\boldsymbol{A} + \boldsymbol{\Delta A})\boldsymbol{x} - \boldsymbol{b}\|_2^2 \; d\boldsymbol{\Delta A} \;\; = \;\; \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \tilde{\rho}\|\boldsymbol{x}\|_2^2,$$

where $\tilde{\rho}$ is a function of the problem dimension n and the uncertainty set size parameter ρ, see Bertsimas and Ma (2023) for more details.

2.2. Properties of the Solution of RLS

Suppose we solve RLS and obtain an optimal solution $\overline{\boldsymbol{x}}$ with optimal value $\bar{c}$. From the definition of the min-max robust problem it follows that $\|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b}\|_2 \leq \bar{c}$, for every matrix $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$, $\boldsymbol{\Delta A} \in \mathcal{U}_2$. Moreover, each matrix $\overline{\boldsymbol{A}}$ satisfies $\|\overline{\boldsymbol{A}} - \boldsymbol{A}\|_F = \|\boldsymbol{A} + \boldsymbol{\Delta A} - \boldsymbol{A}\|_F = \|\boldsymbol{\Delta A}\|_F \leq \rho$. Therefore, we notice that $\overline{\boldsymbol{x}}$ is an approximate solution, with norm error bounded by $\bar{c}$, of infinitely many linear systems $\overline{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$, where each matrix $\overline{\boldsymbol{A}}$ satisfies $\|\overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \leq \rho$. We next illustrate that $\overline{\boldsymbol{x}}$ is also the exact solution of multiple linear systems on nearby matrices.

Theorem 2 (RLS as exact solution on nearby linear systems). *Let* $\overline{\boldsymbol{x}}$ *denote the optimal solution of Problem* (5) *with optimal value* $\bar{c}$*. Then, for each matrix* $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}, \boldsymbol{\Delta A} \in \mathcal{U}_{11}$*, there exists an interval* $\left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$*, with the i-th row of* $\boldsymbol{\Sigma}$ *defined as* $\boldsymbol{\sigma}_i = \bar{c}\frac{|\overline{\boldsymbol{x}}|}{\|\overline{\boldsymbol{x}}\|_2^2}$*, containing a matrix* $\hat{\boldsymbol{A}}$ *such that* $\overline{\boldsymbol{x}}$ *is the exact solution of the linear system* $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$*. Moreover, each matrix* $\hat{\boldsymbol{A}}$ *satisfies* $\|\hat{\boldsymbol{A}} - \boldsymbol{A}\|_F \leq \sqrt{n}\left(\frac{\bar{c}}{\|\overline{\boldsymbol{x}}\|_2} + \rho\right)$.

We refer to the Appendix for a discussion of this result for uncertainty set $\mathcal{U}_2$ and Problem (3).
Moreover, we illustrate that the robust solution is also an approximate solution, with bounded norm error, on a family of preconditioned linear systems, that is linear systems of the form $\boldsymbol{MAx} = \boldsymbol{Mb}$. We have the following result:

Theorem 3 (Preconditioning based on RLS). *Let $\overline{\boldsymbol{x}}$ denote the optimal solution of Problem* (3) *with optimal value $\bar{c}$. Then, for $\theta \in [0, \rho]$, $\overline{\boldsymbol{x}}$ is an approximate solution, with norm error bounded by $\bar{c} + \theta \frac{\|\boldsymbol{b}\|_2}{\|\boldsymbol{A}\|_F}$, of the preconditioned linear system $\boldsymbol{M}(\theta)\boldsymbol{Ax} = \boldsymbol{M}(\theta)\boldsymbol{b}$, where $\boldsymbol{M}(\theta) = (1 + \frac{\theta}{\|\boldsymbol{A}\|_F})\boldsymbol{I}$.*

We note that in this case the bound on the norm error of the robust solution is not as tight as that for the linear systems $\overline{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$.

Connection with Interval Linear Systems

The field of interval linear systems deals with uncertain linear systems, where the uncertainty set is an interval, that is $\boldsymbol{A} \in \left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$. Current results suggest that for a given $\boldsymbol{x}$ checking weak solvability of the interval linear system, that is existence of a matrix $\hat{\boldsymbol{A}} \in \left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$ for which $\boldsymbol{x}$ solves the linear system $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$, as well as strong solvability, that is solving the linear system for every matrix in the interval, are both NP-hard problems (Fiedler et al., 2006). Unless some structure is assumed on the underlying matrix, that is $\boldsymbol{A}^{-1} \geq \boldsymbol{0}$ or $A_{ii} > 0$, $A_{ij} \leq 0, i \neq j$, interval linear systems are computationally intractable Corsaro and Marino (2006). Most algorithms for interval linear systems iteratively compute a tight enclosure of the solution set, that is a vector interval that contains a vector $\boldsymbol{x}$ which solves a linear system for some matrix in the interval, see (Corsaro and Marino, 2006). On the other hand, in this paper we illustrated that there exists an interval around each matrix $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$, $\boldsymbol{\Delta A} \in \mathcal{U}$, such that $\overline{\boldsymbol{x}}$, the optimal solution of RLS, is a weak solution of the corresponding interval linear system.

2.3. Nearly Singular Matrices and the Benefit of RLS

One central result of the paper is the ability of RO to provide more accurate solutions than the classical linear algebra methods in the regime of nearly singular matrices, a class of matrices arising quite often in practice, i.e. covariance matrices coming from data with fewer samples than features.

The main quantity utilized for measuring singularity is the condition number of a matrix which is defined as $\text{cond}(\boldsymbol{A}) = \|\boldsymbol{A}\|_2\|\boldsymbol{A}^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}$, where σ_1 and σ_n denote the largest and smallest singular values of $\boldsymbol{A}$ respectively. The definition of a nearly singular matrix is that of a matrix with large condition number (Datta, 2010; Trefethen and Bau III, 1997). The case $\text{cond}(\boldsymbol{A}) = \infty$ corresponds to a singular matrix. Guaranteed accurate solutions for CLS exist only in the case of well-conditioned matrices. More specifically, if $\boldsymbol{x}$ denotes the true unknown solution, $\hat{\boldsymbol{x}}$ the computed one from CLS and $\tilde{r} = \frac{\|\boldsymbol{A}\hat{\boldsymbol{x}}-\boldsymbol{b}\|_2}{\|\boldsymbol{b}\|_2}$ the relative residual, the following bound holds (Datta, 2010)

$$\frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} \leq \tilde{r}\,\text{cond}(\boldsymbol{A}).$$

We observe that the relative error of the linear system solution is $\mathcal{O}(\tilde{r}\text{cond}(\boldsymbol{A}))$. Considering that in general $\tilde{r}$ is small, the relative error of the solution depends mostly on $\text{cond}(\boldsymbol{A})$. We next illustrate that $\text{cond}(\boldsymbol{A})$ is reduced, when considering RLS with each norm term in the objective squared, leading to a more accurate solution.

Theorem 4 (The effect of robustness on conditioning).
Assume we have a linear system with input data $\boldsymbol{A}, \boldsymbol{b}$. Let σ_1, σ_n denote the largest and smallest singular values of $\boldsymbol{A}^T\boldsymbol{A}$. Then, RLS with each norm term squared is guaranteed to find a solution with relative error $\mathcal{O}\left(\tilde{r}_R \frac{\sigma_1+\rho}{\sigma_n+\rho}\right)$, where $\tilde{r}_R$ is the relative residual of the robust solution, whereas CLS is guaranteed to find one with relative error $\mathcal{O}\left(\tilde{r}_N \frac{\sigma_1}{\sigma_n}\right)$, where $\tilde{r}_N$ is the relative residual of the nominal solution.

Proof. For the nominal linear system we have $\min_{\boldsymbol{x}}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2$ and for the robust with each term in the objective squared we have $\min_{\boldsymbol{x}}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2+\rho\|\boldsymbol{x}\|_2^2$. Computing the gradient and setting it to zero, we obtain $\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{x}}^{NOM} = \boldsymbol{A}^T\boldsymbol{b}$ for the nominal and $(\boldsymbol{A}^T\boldsymbol{A} + \rho\boldsymbol{I})\hat{\boldsymbol{x}}^{ROB} = \boldsymbol{A}^T\boldsymbol{b}$ for the robust. Let $\boldsymbol{x}^{NOM}, \boldsymbol{x}^{ROB}$ denote the exact unknown solutions for the nominal and robust linear systems respectively. Let $\tilde{r}_N$, $\tilde{r}_R$ denote the relative residuals of the nominal and robust solutions respectively. From (Trefethen and Bau III, 1997; Datta, 2010), the relative error of the nominal linear system solution is $\mathcal{O}(\tilde{r}_N\text{cond}(\boldsymbol{A}^T\boldsymbol{A}))$, while that of the robust linear system solution is $\mathcal{O}(\tilde{r}_R\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I}))$. Further, observe that $\sigma_i(\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I}) = \sigma_i(\boldsymbol{A}^T\boldsymbol{A})+\rho$,

thus we obtain

$$\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I}) = \frac{\sigma_1+\rho}{\sigma_n+\rho}.$$

Combining all the above we have the following relative errors

$$\frac{\|\hat{\boldsymbol{x}}^{NOM}-\boldsymbol{x}^{NOM}\|_2}{\|\boldsymbol{x}^{NOM}\|_2} = \mathcal{O}\left(\tilde{r}_N\frac{\sigma_1}{\sigma_n}\right),$$

$$\frac{\|\hat{\boldsymbol{x}}^{ROB}-\boldsymbol{x}^{ROB}\|_2}{\|\boldsymbol{x}^{ROB}\|_2} = \mathcal{O}\left(\tilde{r}_R\frac{\sigma_1+\rho}{\sigma_n+\rho}\right). \quad \square$$

We can think of the uncertainty set size parameter ρ as a way of improving the conditioning of the matrix $\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I}$, thus resulting in a more accurate solution, assuming a small relative residual. In Table 1 we illustrate how the condition number of the matrix $\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I}$ is reduced when increasing ρ.

Table 1: The effect of robustness on the condition number of $\boldsymbol{A}^T\boldsymbol{A}$ for a randomly generated ill-conditioned matrix.

$\text{cond}(\boldsymbol{A}^T\boldsymbol{A})$	$\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+0.01\boldsymbol{I})$	$\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+0.1\boldsymbol{I})$	$\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+0.5\boldsymbol{I})$
10^{16}	10^4	10^3	10^2
10^{17}	10^4	10^3	10^2
10^{18}	10^7	10^6	10^5

From Table 1 one might conclude that we can keep increasing ρ to further reduce $\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I})$ and thus obtain a more accurate solution of the linear system. However, we note that as we increase ρ there is less weight placed on minimizing the norm error of the solution, which can result in solutions with large relative residuals. In the regime of nearly singular matrices, for small values of ρ the term $\text{cond}(\boldsymbol{A}^T\boldsymbol{A}+\rho\boldsymbol{I})$ drives the accuracy of the solution, however for larger values of ρ the norm error of the obtained robust solution will increase, leading to a large relative residual and therefore an inaccurate solution.

2.4. Numerical Experiments

In this section, we provide computational results on RLS, in the case of nearly singular matrices. We compare RLS, with Gaussian elimination as well as the current state of the art iterative methods for linear systems. The

data generation process is as follows. First, we generate a $(n-1) \times n$ matrix with entries from the interval $[0, 1]$ and then add a row that is the sum of previous ones with small added noise. The noise vector is generated from $\mathcal{N}(0, \sigma^2)$ where σ controls the degree of singularity of the matrix. Matrices with condition numbers 10^9 and 10^{17} can be generated with $\sigma \approx 10^{-5}$ and $\sigma \approx 10^{-13}$ respectively. The right-hand side vector $\boldsymbol{b}$ is generated from $\mathcal{N}(0, 1)$. Throughout the experiments, the parameter ρ is fixed at 0.1.
For each size of input matrix, we report the average error over 1000 randomly generated nearly singular matrices along with the standard deviation for the classical linear system with Gaussian elimination, the RLS and two current state of the art iterative algorithms MINRES, GMRES. The results are illustrated in Tables 2, 3 when the calculations are performed under single and double precision respectively.

Table 2: Results for linear systems on nearly singular matrices, under single precision. Each entry in the columns CLS, RLS, MINRES, GMRES corresponds to the average norm error $\|\boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{b}\|_2$, where $\boldsymbol{x}^*$ is the optimal solution obtained from the corresponding method, over 1000 randomly generated matrices, along with the standard deviation.

$\boldsymbol{n}$	**cond(A)**	**CLS**	**RLS**	**MINRES**	**GMRES**
10^2	10^9	17.78 ± 68.99	$\mathbf{1.43 \pm 0.75}$	16.41 ± 6.31	8.82 ± 0.75
10^3	10^9	299.95 ± 105.98	$\mathbf{2.45 \pm 0.78}$	78.33 ± 67.42	31.35 ± 0.73
10^4	10^9	459.88 ± 501.21	$\mathbf{4.48 \pm 0.82}$	109.75 ± 29.54	99.95 ± 0.91

Table 3: Results for linear systems on nearly singular matrices, under double precision. Each entry in the columns CLS, RLS, MINRES, GMRES corresponds to the average norm error $\|\boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{b}\|_2$, where $\boldsymbol{x}^*$ is the optimal solution obtained from the corresponding method, over 1000 randomly generated matrices, along with the standard deviation.

$\boldsymbol{n}$	**cond(A)**	**CLS**	**RLS**	**MINRES**	**GMRES**
10^2	10^{17}	4.43 ± 45.67	$\mathbf{1.49 \pm 0.75}$	16.67 ± 6.41	8.85 ± 0.73
10^3	10^{18}	94.22 ± 242.12	$\mathbf{2.45 \pm 0.79}$	71.12 ± 60.82	31.35 ± 0.71
10^4	10^{18}	142.55 ± 356.77	$\mathbf{3.85 \pm 0.82}$	109.88 ± 45.55	99.91 ± 1.05

From Tables 2, 3, we observe that RLS achieves the smallest norm error in all cases. Moreover, we notice that in CLS the standard deviation is very high because of solutions with very large norm error, as a result of applying Gaussian elimination on ill-conditioned matrices.

One important aspect is the separation of round-off errors and data uncertainty. It is well known that Gaussian elimination does not behave well in ill-conditioned problems, since $\log_{10} \text{cond}(\boldsymbol{A})$ digits are lost in calculations (Trefethen and Bau III, 1997). Therefore, assuming precision s and a matrix $\boldsymbol{A}$ with $\text{cond}(\boldsymbol{A}) = 10^t$, the number of digits lost in calculations are $s - t$. As a result, the threshold for RLS to do better than CLS is when $t \geq s$, which is the regime of nearly singular matrices. In single precision it corresponds to $t \geq 7$ and in double precision to $t \geq 16$. Figure 1 illustrates this behavior.

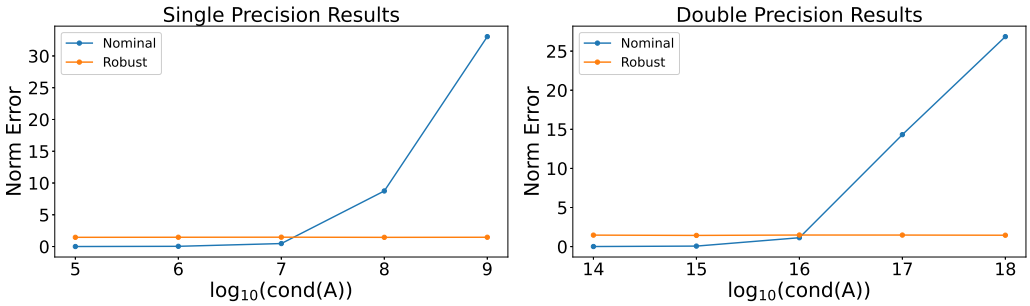


Figure 1: Behavior of nominal and robust linear system error under single and double precision

In Table 4, we provide a comparison of the different approaches for RLS, that is Problems (3), (4) and (5). We notice that in all cases, Problem (3) achieves the lowest norm error, with the difference between the other two being more pronounced for $n = 10^3$ and significantly more pronounced for $n = 10^4$. Moreover, we notice that Problem (5) constantly achieves slightly better norm error than Problem (4).

Table 4: Comparisons of different methods for RLS under double precision. Each entry in the columns RLS-22, RLS-21, RLS-11 corresponds to the average norm error of RLS solved with either Problem (3), (4) or (5) respectively, over 1000 randomly generated matrices, along with the standard deviation.

n	**RLS-11**	**RLS-21**	**RLS-22**
10^2	4.08 ± 1.42	4.12 ± 0.79	$\mathbf{1.49 \pm 0.75}$
10^3	10.95 ± 2.26	12.87 ± 0.84	$\mathbf{2.45 \pm 0.79}$
10^4	35.55 ± 1.12	40.18 ± 0.95	$\mathbf{3.85 \pm 0.82}$

Further, in Table 5 we report computational times of all different methods

for linear systems. We observe that the nominal approach requires the least computational time across all dimensions and that among the different robust approaches, the one corresponding to Problem (3) is always solved faster. Further, we notice that for $n = 10^2$, all methods require on average less than a second to solve a linear system. Moreover, for $n = 10^3$ the nominal approach as well as the iterative methods need on average less than a second, while the robust ones need on average $1.66-9.72$ seconds. Finally, for $n = 10^4$ the nominal approach requires the least amount of computational time, that is 3 seconds on average. Moreover, we notice a significant difference between the different robust approaches, where the one corresponding to Problem (3) needs on average 526.53 seconds, while those corresponding to Problems (4) and (5) need on average 1880.27 and 2551.78 seconds respectively.

Table 5: Comparisons of computational times between different methods for linear systems . Columns RLS-22, RLS-21, RLS-11 refer to RLS solved with either Problem (3), (4) or (5) respectively. All reported times are in seconds, averaged over 1000 instances.

n	CLS	RLS-11	RLS-21	RLS-22	MINRES	GMRES
10^2	0.01 ± 0.01	0.04 ± 0.07	0.21 ± 0.09	0.03 ± 0.09	0.01 ± 0.03	0.01 ± 0.08
10^3	0.02 ± 0.01	9.72 ± 0.87	4.49 ± 0.45	1.66 ± 0.22	0.10 ± 0.01	0.11 ± 0.01
10^4	3.05 ± 0.49	2551.78 ± 16.54	1880.27 ± 50.19	526.53 ± 37.77	508.68 ± 5.01	534.46 ± 4.89

Finally, we compare the performance of RLS with that of state of the art iterative methods on sparse nearly singular matrices. For a $(n-1) \times n$ randomly generated matrix from $\mathcal{N}(0,\, 1)$ we randomly sample some indices and assign them to zero. Then, we add a row that is the sum of the previous ones with small added noise, so that the resulting matrix is nearly singular. As Figure 2 illustrates, when the matrix becomes more sparse while remaining nearly singular, the norm error of all methods increases. We notice that RLS constantly achieves the smallest norm error. More precisely, for both $n = 10^2$ and $n = 10^3$ the improvement in norm error of RLS over the other methods is more pronounced on non-sparse matrices, i.e., matrices with less than 25% zero entries. When then matrix has between 25% and 50% zero entries, RLS achieves the smallest norm error but with a smaller difference than GMRES. Finally, for $n = 10^3$ we observe that if the matrix has more than 75% zero entries both RLS and GMRES achieve similar norm errors, with the former being slightly better.

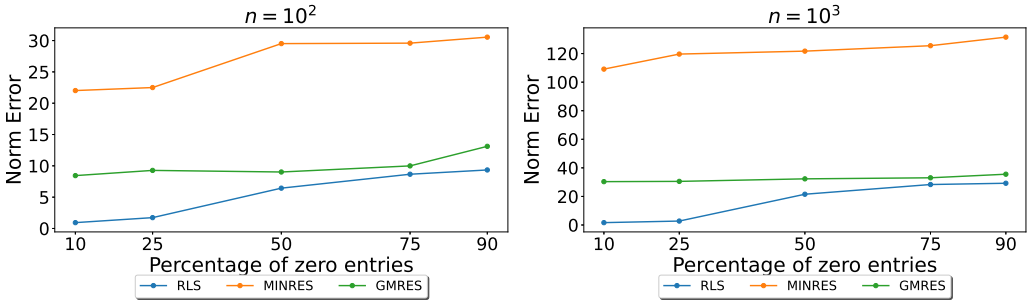


Figure 2: The effect of sparsity of the input matrix on the norm error of the linear system. The x-axis represents the percentage of zeros values of the matrix $\boldsymbol{A}$ and the y-axis the norm error of RLS solved with Problem (3), MINRES and GMRES.

3. Robust Matrix Inversion

In this section, we address matrix inversion under uncertainty. Leveraging the methodology introduced for RLS, we develop a method for calculating a robust matrix inverse (RMI). In the numerical experiments, we show the benefit of our approach over state of the art methods for matrix inversion, while focusing on nearly singular matrices.

3.1. Formulation

We present the classical approach to compute the inverse of a matrix $\boldsymbol{A}$. one has to do pivots (row operations) in order to create the identity matrix out of a given matrix. The classical procedure is as follows: Start with the matrix $\boldsymbol{A}$, then horizontally stack the identity next to $\boldsymbol{A}$ and get the matrix $[\boldsymbol{A}|\boldsymbol{I}]$ and do pivoting in order to end up in the form $[\boldsymbol{I}|\boldsymbol{B}]$.Then, we have $\boldsymbol{A}^{-1} = \boldsymbol{B}$.
Matrix inversion can also be viewed as a sequence of n linear systems

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{e_i}, \;\; i \in [n],$$

where the solution of the i-th linear system is the i-th column of the inverse. We propose minimizing the norm error, that is

$$\min_{\boldsymbol{x}} \;\; \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{e_i}\|_2, \;\; i \in [n]. \tag{6}$$

Following the model of uncertainty we introduced for RLS, the robust problem in order to compute the i-th column of RMI is formulated as follows:

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{\Delta A} \in \mathcal{U}} \|(\boldsymbol{A} + \boldsymbol{\Delta A})\boldsymbol{x} - \boldsymbol{e}_i\|_2. \tag{7}$$

The RC of Problem (7) is as follows (see Section 2.1):

$$\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{e}_i\|_2 + \rho\|\boldsymbol{x}\|_2. \tag{8}$$

Similarly, we can compute the RC for uncertainty set $\mathcal{U}_1$ and $\mathcal{U}_{11}$, see Section 2.1.

3.2. Algorithms for Matrix Inversion under Uncertainty

We can compute a RMI by solving a RLS for each column. Moreover, in order to speed up computational time, we can solve the linear systems in parallel. We summarize our approach in pseudocode in Algorithm 1.

Algorithm 1 Robust Inverse Computation

Input: Matrix $\boldsymbol{A}$, Parameter ρ.
Output: Matrix $\boldsymbol{B}$, the robust inverse of $\boldsymbol{A}$.

1: Initialize $\boldsymbol{B}$ as an empty array.
2: **for** $k = 1 : n$ **do**
3: $\quad \boldsymbol{y} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{e}_k\|_2 + \rho\|\boldsymbol{x}\|_2$.
4: $\quad$ Append $\boldsymbol{y}$ to $\boldsymbol{B}$.
5: **end for**
6: Return $\boldsymbol{B}$.

We note that since RLS are mostly advantageous on nearly singular matrices, we expect RMI to follow the same behavior.

3.3. Inverse Covariance Estimation

A special case of nearly singular matrices that arise often are covariance matrices from data. For a data matrix $\boldsymbol{X} \in \mathbb{R}^{n\times p}$ the covariance matrix is defined as $\boldsymbol{\Sigma} = \frac{\boldsymbol{X}^T\boldsymbol{X}}{n-1}$. When $n < p$, this matrix is nearly singular and in the case when $n << p$ it becomes singular. One problem of interest in many applications is inverse covariance estimation (Friedman et al., 2008). For that purpose, the following optimization problem has been proposed (Yuan and Lin, 2007)

$$\min_{\boldsymbol{\Theta}} \quad \operatorname{tr}(\boldsymbol{\Sigma}^T\boldsymbol{\Theta}) - \log\det\boldsymbol{\Theta} + \rho\|\boldsymbol{\Theta}\|_1. \tag{9}$$

The objective of Problem (9) originates from the negative log-likelihood of the data (Yuan and Lin, 2007). Moreover, the optimal solution $\boldsymbol{\Theta}^*$ is a sparse inverse covariance estimation. Problem (9) is also known as GLasso and has

been introduced to induce sparsity. However, it can also be characterized as the RC of the following problem:

$$\min_{\boldsymbol{\Theta}} \max_{\boldsymbol{\Delta\Sigma}\in\mathcal{V}} < \boldsymbol{\Sigma} + \boldsymbol{\Delta\Sigma}, \boldsymbol{\Theta} > - \log\det(\boldsymbol{\Theta}), \tag{10}$$

where $\mathcal{V} = \{\boldsymbol{X} \in \mathbb{R}^{p\times p} : \|\boldsymbol{X}_i\|_\infty \leq \rho\}$.

Theorem 5 (GLasso equivalence to robustness). *The RC of Problem* (10) *is Problem* (9)*.*

Proof. The objective of Problem (10), can be written as

$$\langle \boldsymbol{\Sigma} + \boldsymbol{\Delta\Sigma}, \boldsymbol{\Theta}\rangle - \log\det(\boldsymbol{\Theta}) = \langle \boldsymbol{\Sigma}, \boldsymbol{\Theta}\rangle + \langle \boldsymbol{\Delta\Sigma}, \boldsymbol{\Theta}\rangle - \log\det(\boldsymbol{\Theta})$$

$$\langle \boldsymbol{\Delta\Sigma}, \boldsymbol{\Theta}\rangle = \mathrm{tr}(\boldsymbol{\Delta\Sigma}^T\boldsymbol{\Theta}) = \sum_{i=1}^{p} \boldsymbol{\Delta\Sigma}_i^T\boldsymbol{\Theta}_i \leq \sum_{i=1}^{p} \rho\|\boldsymbol{\Theta}_i\|_1 = \rho\|\boldsymbol{\Theta}\|_1$$

For $\boldsymbol{\Delta\Sigma}_{ij} = \rho\ sign(\Theta_{ij})$ we have

$$\langle \boldsymbol{\Delta\Sigma}, \boldsymbol{\Theta}\rangle = \sum_i \sum_j \Delta\Sigma_{ij}\Theta_{ij} = \sum_i \sum_j \rho\, sign(\Theta_{ij})\Theta_{ij} = \sum_i \rho\|\boldsymbol{\Theta}_i\|_1 = \rho\|\boldsymbol{\Theta}\|_1$$

$$\implies \max_{\boldsymbol{\Delta\Sigma}\in V} \langle \boldsymbol{\Sigma} + \boldsymbol{\Delta\Sigma}, \Theta\rangle - \log\det(\boldsymbol{\Theta}) = \langle \boldsymbol{\Sigma}, \boldsymbol{\Theta}\rangle - \log\det(\boldsymbol{\Theta}) + \rho\|\boldsymbol{\Theta}\|_1. \quad \square$$

3.4. Numerical Experiments

In this section, we compare RMI with state of the art approaches for matrix inversion. We generate random matrices with entries from $[0, 1]$ and add a row that is the sum of previous ones with small added noise so that the resulting matrix is nearly singular. We report the norm error, that is $\|\boldsymbol{A}\boldsymbol{X}^* - \boldsymbol{I}\|_F$, where $\boldsymbol{X}^*$ is the obtained inverse. The results are illustrated in Table 6.

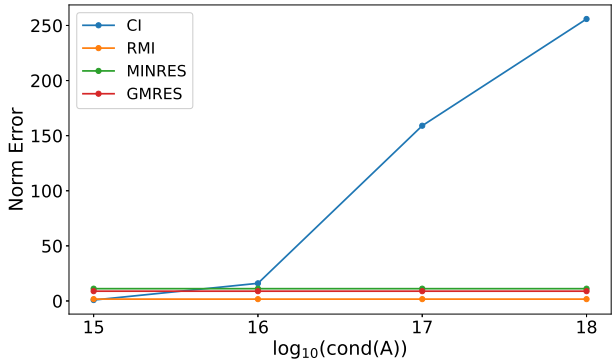
As Table 6 illustrates, RMI outperforms the other approaches in terms of norm error. We observe that the difference is more significant for $n = 10^3$. We next examine the behavior of the errors of the robust and nominal matrix inversion with the degree of singularity of the input matrix. The results are illustrated in Figure 3.

From Figure 3 we observe that if $\mathrm{cond}(\boldsymbol{A}) \leq 10^{15}$, the nominal inverse achieves the smallest norm error. However, if $\mathrm{cond}(\boldsymbol{A}) \geq 10^{16}$ then the robust inverse achieves the smallest norm error. Moreover, we notice that

Table 6: Results for matrix inversion on nearly singular matrices. Columns CI, RMI, MINRES, GMRES correspond to the norm error of the method averaged over 100 randomly generated matrices, along with the standard deviation.

n	cond(A)	CI	RMI	MINRES	GMRES
10^2	10^{17}	461.81 ± 195.02	$\mathbf{2.91 \pm 0.15}$	10.41 ± 0.17	8.58 ± 0.19
10^3	10^{18}	3183.55 ± 800.07	$\mathbf{4.26 \pm 0.1}$	32.83 ± 0.49	31.09 ± 0.53

Figure 3: Matrix inversion norm error as the matrix approaches singularity for nominal and robust. The x-axis represents cond($\boldsymbol{A}$) in a $\log_{10}$ scale and the y-axis represents the norm error for nominal(CI), robust(RMI), MINRES and GMRES.



the nominal inverse exhibits significant norm error, as the matrix approaches singularity, an observation aligned with Figure 1, since the matrix inverse carries out the norm errors of multiple linear systems. Finally, we observe that solving each linear system with the robust approach instead of MINRES or GMRES results in a more accurate inverse independent of cond($\boldsymbol{A}$).

In Table 7 we present computational times of different methods for matrix inversion. We observe that the nominal approach requires the least computational time in all cases. Moreover, the robust approach with problem (3) as well as the iterative methods require on average less than a second for $n = 10^2$. Further, for $n = 10^3$ MINRES and GMRES require on average 126.12 and 143.08 seconds respectively, while the robust approach with problems (3) and (4) require on average 872.61 and 2318.99 seconds respectively. Finally, we compare RMI with GLasso, while focusing on nearly singular matrices. We generated synthetic covariance matrices that are nearly singular and compute RMI with Algorithm 1 and GLasso implemented in the R package GLasso (Friedman et al., 2015). We varied $\rho \in [0.011]$ and reported

Table 7: Computational times of different methods for matrix inversion. Columns RMI-22 and RMI-21 refer to RMI with each RLS solved with either problem (3) or (4) respectively. All reported times are in seconds, averaged over 100 instances.

n	CI	RLS-21	RLS-22	MINRES	GMRES
10^2	0.01 ± 0.003	1.49 ± 2.09	0.67 ± 0.16	0.15 ± 0.05	0.16 ± 0.11
10^3	0.04 ± 0.02	2318.99 ± 35.62	872.61 ± 22.25	126.12 ± 19.98	143.08 ± 8.89

the best norm error for each method.

Table 8: Norm error comparison between RMI, CI and GLasso, averaged over 100 covariance matrices. Column (n, p) refers to the dimensions of matrix $\boldsymbol{X}$. The p-value corresponds to the null hypothesis of equal norm errors between RMI and Glasso.

(n, p)	cond(A)	CI	RMI	GLasso	p-value
$(100, 105)$	10^{17}	98.31	**2.38**	2.85	10^{-3}
$(500, 550)$	10^{18}	521.12	**7.08**	9.08	$< 10^{-5}$
$(1500, 1600)$	10^{18}	3562.11	**10.42**	18.55	$< 10^{-5}$

As Table 8 illustrates, RMI achieves smaller norm error than GLasso across the board, with the difference being more significant for $(n, p) = (1500, 1600)$. Further, we notice that the difference in average norm errors between RMI and GLasso is statistically significant in all instances, as indicated with the p-values. This observation follows from the fact that the uncertainty set $\mathcal{V}$, corresponding to GLasso is more conservative than the uncertainty set $\mathcal{U}_2$, corresponding to RMI. For example, the matrix $\boldsymbol{\Delta A}$ with $\Delta A_{ij} = \rho$ is contained in $\mathcal{V}$ but not in $\mathcal{U}_2$.

4. Robust Eigenvalues-Eigenvectors

In this section, we derive an algorithm for computing eigen-pairs under uncertainty, utilizing alternating optimization and the results from Section 2. We illustrate the advantages over the nominal eigen-pairs, with numerical experiments on synthetic nearly singular matrices.

4.1. Formulation

For a symmetric matrix $\boldsymbol{A}$ the eigenvalue problem consists of finding n scalars $\lambda_i, i \in [n]$ and n vectors $\boldsymbol{x}_i, i \in [n]$ that satisfy the following:

$$\begin{aligned} &\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i, \ \ \forall i \in [n], \\ &\boldsymbol{x}_i^T \boldsymbol{x}_j = 0, \ \ \forall i \neq j, \\ &||\boldsymbol{x}_i||_2 = 1, \ \ \forall i \in [n]. \end{aligned}$$

Instead of satisfying the first equality constraint exactly, we want to compute a pair $(\lambda, \boldsymbol{x})$ that minimizes the norm error, that is

$$\min_{\lambda, \boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{x}\|_2. \tag{11}$$

In case of uncertainty in the entries of $\boldsymbol{A}$ we obtain the following problem:

$$\min_{\lambda, \boldsymbol{x}} \max_{\boldsymbol{\Delta A} \in \mathcal{U}} \|(\boldsymbol{A} + \boldsymbol{\Delta A})\boldsymbol{x} - \lambda \boldsymbol{x}\|_2, \tag{12}$$

where we solve for the worst possible realization of uncertainty. The RC for $\mathcal{U} = \mathcal{U}_2$ (see Section 2.1), is as follows:

$$\min_{\lambda, \boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{x}\|_2 + \rho \|\boldsymbol{x}\|_2. \tag{13}$$

Similarly, we can derive the RC for $\mathcal{U} = \mathcal{U}_1$ and $\mathcal{U} = \mathcal{U}_{11}$, see Section 2.1.

Observe that for a known eigenvalue we can compute the robust eigenvector by solving Problem (13). If we know all n eigenvalues, in order to compute all n robust eigenvectors, we have to solve n such problems, while also requiring that the eigenvectors are orthonormal. Assuming we have computed the eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}$, in the i-th eigenvector problem we add the linear constraints $\boldsymbol{x}^T \boldsymbol{x}_j = 0, \ \forall j \in [i-1]$ to ensure orthogonality and normalize the optimal solution in order to have unit norm. The problem corresponding to the i-th eigenvector is formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\boldsymbol{A}\boldsymbol{x} - \lambda_i \boldsymbol{x}\|_2 + \rho \|\boldsymbol{x}\|_2 \\ \text{s.t.} \quad & \boldsymbol{x}^T \boldsymbol{x}_j = 0, \ \ j \in [i-1]. \end{aligned} \tag{14}$$

Assuming that we know all n eigenvectors, we can then compute the eigenvalues by solving each problem over λ_i. The problem for the i-th eigenvalue is as follows:

$$\min_{\lambda_i} \|\boldsymbol{A}\boldsymbol{x}_i - \lambda_i \boldsymbol{x}_i\|_2. \tag{15}$$

Observe that Problem (15) has the closed form solution

$$\lambda_i^* = \frac{\boldsymbol{x}_i^T \boldsymbol{A} \boldsymbol{x}_i}{\boldsymbol{x}_i^T \boldsymbol{x}_i} \;=\; \boldsymbol{x}_i^T \boldsymbol{A} \boldsymbol{x}_i.$$

We next illustrate that when computing robust eigenvectors the condition number of the underlying matrix can be improved over the nominal.

Theorem 6 (Robustness effect on conditioning for eigen-pairs). *Assume input matrix $\boldsymbol{A} \in \mathcal{S}_+^n$ with ordered eigenvalues $(\lambda_1, \ldots, \lambda_n)$. Then, for a fixed eigenvalue λ, an upper bound for the condition number of the underlying matrix for computing the robust eigenvector with each term in the objective squared, is $\frac{|\lambda_1^2 - 2\lambda\lambda_n + \lambda^2 + \rho|}{|\lambda_n^2 - 2\lambda\lambda_1 + \lambda^2 + \rho|}$, while for that of the nominal eigenvector, is $\frac{|\lambda_1^2 - 2\lambda\lambda_n + \lambda^2|}{|\lambda_n^2 - 2\lambda\lambda_1 + \lambda^2|}$.*

Proof. For the nominal eigenvector we have $\min_{\boldsymbol{x}} \frac{1}{2}\|(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}\|_2^2$, while for the robust we have $\min_{\boldsymbol{x}} \frac{1}{2}\|(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}\|_2^2 + \frac{1}{2}\rho\|\boldsymbol{x}\|_2^2$. Computing the gradient and setting it to zero, we obtain $(\boldsymbol{A} - \lambda \boldsymbol{I})^T(\boldsymbol{A} - \lambda \boldsymbol{I})\hat{\boldsymbol{x}}^{NOM} = \boldsymbol{0}$ for the nominal and $((\boldsymbol{A} - \lambda \boldsymbol{I})^T(\boldsymbol{A} - \lambda \boldsymbol{I}) + \rho \boldsymbol{I})\hat{\boldsymbol{x}}^{ROB} = \boldsymbol{0}$ for the robust. We have $(\boldsymbol{A} - \lambda \boldsymbol{I})^T(\boldsymbol{A} - \lambda \boldsymbol{I}) = \boldsymbol{A}^T\boldsymbol{A} - 2\lambda \boldsymbol{A} + \lambda^2 \boldsymbol{I}$. Further, the ordered eigenvalues of $\boldsymbol{A}^T\boldsymbol{A} + \lambda^2 \boldsymbol{I}$ are $(\lambda_1^2 + \lambda^2, \ldots, \lambda_n^2 + \lambda^2)$. From Weyl's inequality (Weyl, 1912) it follows that

$$\frac{\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A} + \lambda^2 \boldsymbol{I} - 2\lambda \boldsymbol{A})}{\lambda_{min}(\boldsymbol{A}^T\boldsymbol{A} + \lambda^2 \boldsymbol{I} - 2\lambda \boldsymbol{A})} \;\leq\; \frac{\lambda_1^2 + \lambda^2 - 2\lambda\lambda_n}{\lambda_n^2 + \lambda^2 - 2\lambda\lambda_1}.$$

Therefore, in the worst-case for the nominal we have

$$\mathrm{cond}(\boldsymbol{A}^T\boldsymbol{A} - 2\lambda \boldsymbol{A} + \lambda^2 \boldsymbol{I}) = \frac{|\lambda_1^2 - 2\lambda\lambda_n + \lambda^2|}{|\lambda_n^2 - 2\lambda\lambda_1 + \lambda^2|},$$

and for the robust we have

$$\mathrm{cond}(\boldsymbol{A}^T\boldsymbol{A} - 2\lambda \boldsymbol{A} + \lambda^2 \boldsymbol{I} + \rho \boldsymbol{I}) = \frac{|\lambda_1^2 - 2\lambda\lambda_n + \lambda^2 + \rho|}{|\lambda_n^2 - 2\lambda\lambda_1 + \lambda^2 + \rho|}. \qquad \square$$

We note that for $\lambda \approx 0$ and $\rho \neq 0$, the underlying matrix for the robust eigenvector ($\boldsymbol{A}^{ROB}$) is better conditioned than the one for the nominal ($\boldsymbol{A}^{NOM}$), that is,

$$\mathrm{cond}(\boldsymbol{A}^{NOM}) = \frac{\lambda_1^2}{\lambda_n^2}, \;\; \mathrm{cond}(\boldsymbol{A}^{ROB}) = \frac{\lambda_1^2 + \rho}{\lambda_n^2 + \rho}.$$

From Table 9 we observe that matrices with condition numbers 10^{17} or higher, have a significant percentage of almost zero eigenvalues. We note that as illustrated in the proof of Theorem 6, the underlying linear system for both nominal and robust eigen-pairs is homogeneous, thus if we make ρ large enough to obtain a very well conditioned matrix, the unique optimal solution is the trivial solution $\boldsymbol{x} = \mathbf{0}$.

Table 9: Percentage of eigenvalues of $\boldsymbol{A} = \boldsymbol{A}_0^T \boldsymbol{A}_0$ that are below 10^{-5}, i.e., $\lambda(\boldsymbol{A}) \approx 0$. We generate $\boldsymbol{A}_0 \in \mathbb{R}^{(n+k)\times n}$ from $\mathcal{N}(0,1)$ and vary $k \in [-5, 1]$. We use $n = 10$.

$\text{cond}(\boldsymbol{A})$	$\lambda(\boldsymbol{A}) \approx 0$
10^4	0%
10^{15}	10%
10^{17}	60%
10^{18}	80%

An adjustment

We observe that a trivial optimal solution for Problem (14) is $\boldsymbol{x} = \mathbf{0}$. In order to avoid this, we require that $|\boldsymbol{e}^T\boldsymbol{x}| \geq \delta$ for some small $\delta > 0$. Instead of adding the non-convex constraint, we can solve problem

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\boldsymbol{A}\boldsymbol{x} - \lambda_i \boldsymbol{x}\|_2 + \rho\|\boldsymbol{x}\|_2 \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{x}_j = 0, \;\; j \in [i-1], \\ & \boldsymbol{e}^T\boldsymbol{x} \geq \delta, \end{aligned} \tag{16}$$

and

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\boldsymbol{A}\boldsymbol{x} - \lambda_i \boldsymbol{x}\|_2 + \rho\|\boldsymbol{x}\|_2 \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{x}_j = 0, \;\; j \in [i-1], \\ & \boldsymbol{e}^T\boldsymbol{x} \leq -\delta. \end{aligned} \tag{17}$$

Let $\boldsymbol{x}_+^*, \boldsymbol{x}_-^*$ denote the optimal solutions of Problems (16) and (17), respectively. Then, we can keep the one with the smallest norm error. Namely, if $\|\boldsymbol{A}\boldsymbol{x}_+^* - \lambda_i \boldsymbol{x}_+^*\|_2 \leq \|\boldsymbol{A}\boldsymbol{x}_-^* - \lambda_i \boldsymbol{x}_-^*\|_2$ we take $\boldsymbol{x}_i = \boldsymbol{x}_+^*$, else $\boldsymbol{x}_i = \boldsymbol{x}_-^*$. If one of the two problems is infeasible, we can keep the solution of the other one and if both are infeasible we can use a smaller value for δ.
So far we have seen how we can compute the robust eigenvectors for known eigenvalues and the eigenvalues for known eigenvectors. This motivates the

use of an alternating algorithm for computing robust eigen-pairs, where at each step we fix all eigenvalues λ_i and compute the eigenvectors $\boldsymbol{x}_i$ and then the other way around. More precisely, we decompose the problem of finding all robust eigen-pairs as an outer problem over $\boldsymbol{\lambda}$, the n-dimensional vector of eigenvalues and n inner problems, one for each eigenvector $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Thus, one iteration of the alternating algorithm consists of an update on $n+1$ vectors $(\boldsymbol{\lambda}^{(t)}, \boldsymbol{x}_1^{(t)}, \ldots, \boldsymbol{x}_n^{(t)}) \rightarrow (\boldsymbol{\lambda}^{(t+1)}, \boldsymbol{x}_1^{(t+1)}, \ldots, \boldsymbol{x}_n^{(t+1)})$. We can repeat the process until the norm errors $\|\boldsymbol{A}\boldsymbol{x}_i - \lambda_i \boldsymbol{x}_i\|_2$ are very small or stop improving. The overall procedure is summarized in Algorithm 2. The inputs are $\boldsymbol{A}$, the nominal matrix, $\boldsymbol{\lambda}^0$, the vector of initial eigenvalues, T, the number of iterations, ρ the uncertainty set size parameter and δ a small positive number. We can initialize Algorithm 2 with the nominal eigenvalues. Regarding

Algorithm 2 Alternating algorithm for robust eigen-pairs computation.

Input: Matrix $\boldsymbol{A}$, vector $\boldsymbol{\lambda}^0$, parameters T, ρ, δ.
Output: $\boldsymbol{X}$ matrix of robust eigenvectors,$\boldsymbol{\lambda}$ vector of robust eigenvalues.

$\boldsymbol{\lambda} = \boldsymbol{\lambda}^0$.
for $t = 1 : T$ **do**
  Initialize $\boldsymbol{X}^{(t)}$ as an empty array.
  for $j = 1 : n$ **do**
    Solve Problems (16) and (17) with inputs $\lambda_j^{(t)}, \boldsymbol{X}^{(t)}, \delta, \rho$
      and obtain $\boldsymbol{x}_+^*, \boldsymbol{x}_-^*$.
    if $\|\boldsymbol{A}\boldsymbol{x}_+^* - \lambda_i \boldsymbol{x}_+^*\| \leq \|\boldsymbol{A}\boldsymbol{x}_-^* - \lambda_i \boldsymbol{x}_-^*\|$ **then**
      $\boldsymbol{x}_j^{(t)} = \boldsymbol{x}_+^*$.
    else
      $\boldsymbol{x}_j^{(t)} = \boldsymbol{x}_-^*$.
    end if
    $\boldsymbol{x}_j^{(t)} = \frac{\boldsymbol{x}_j^{(t)}}{\|\boldsymbol{x}_j^{(t)}\|_2}$.
    Append $\boldsymbol{x}_j^{(t)}$ in $\boldsymbol{X}^{(t)}$.
    $\lambda_j^{(t)} = {\boldsymbol{x}_j^{(t)}}^T \boldsymbol{A}\boldsymbol{x}_j^{(t)}$.
  end for
end for
Return $\boldsymbol{X}^{(T)}, \boldsymbol{\lambda}^{(T)}$.

convergence of Algorithm 2, we have the following result:

Theorem 7 (Convergence of Algorithm 2). *Algorithm 2 converges, with*

rate sub-linear in the number of iterations.

PROOF. Algorithm 2 is a special case of alternating minimization over the variables $\boldsymbol{\lambda}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. The sub-problem for each block of variables is convex and therefore the algorithm converges with rate $\mathcal{O}(\frac{1}{T})$, where T denotes the number of iterations (Sun and Hong, 2015), (Hong et al., 2017). □

4.2. Properties of Robust Eigen-Pairs

Suppose we apply Algorithm 2 and obtain optimal solutions $\overline{\boldsymbol{\lambda}}, \overline{\boldsymbol{x}}_1, \ldots, \overline{\boldsymbol{x}}_n$, with optimal value for the i-th problem $\bar{c}_i$. Then, the vectors $\overline{\boldsymbol{x}}_i$ are orthogonal to each other and also have unit norm. Moreover, from the definition of the min-max robust problem it follows that $\|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}}_i - \overline{\lambda}_i\overline{\boldsymbol{x}}_i\|_2 \leq \bar{c}_i$, for every matrix $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$, $\boldsymbol{\Delta A} \in \mathcal{U}_2$. This suggests that the robust eigen-pairs are approximate eigen-pairs on nearby matrices $\overline{\boldsymbol{A}}$, satisfying $\|\overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \leq \rho$, see Section 2.2, with norm error for the i-th eigen-pair bounded by $\bar{c}_i$. Moreover, we next illustrate that for the i-th eigen-pair there exists an interval around each matrix $\overline{\boldsymbol{A}}$ containing a matrix $\hat{\boldsymbol{A}}_i$, such that $\hat{\boldsymbol{A}}_i\overline{\boldsymbol{x}}_i = \overline{\lambda}_i\overline{\boldsymbol{x}}_i$.

Theorem 8 (Robust eigen-pairs as exact on nearby matrices). *Let $(\overline{\lambda}_i, \overline{\boldsymbol{x}}_i)$ denote the i-th robust eigen-pair obtained from Algorithm 2 with uncertainty set $\mathcal{U}_1$, with optimal value $\bar{c}_i$. Then, for each matrix $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$, $\boldsymbol{\Delta A} \in \mathcal{U}_{11}$, there exists an interval $\left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$, with the i-th row of $\boldsymbol{\Sigma}$ defined as $\boldsymbol{\sigma}_i = \bar{c}_i|\overline{\boldsymbol{x}}_i|$, containing a matrix $\hat{\boldsymbol{A}}_i$ such that $\hat{\boldsymbol{A}}_i\overline{\boldsymbol{x}}_i = \overline{\lambda}_i\overline{\boldsymbol{x}}_i$. Moreover, the matrix $\hat{\boldsymbol{A}}_i$ satisfies $\|\hat{\boldsymbol{A}}_i - \boldsymbol{A}\|_F \leq \sqrt{n}\,(\bar{c}_i + \rho)$.*

We note that if we instead use the ℓ_2 norm in the objective, along with the uncertainty set $\mathcal{U}_2$ we would obtain a larger interval, see Appendix for more details.

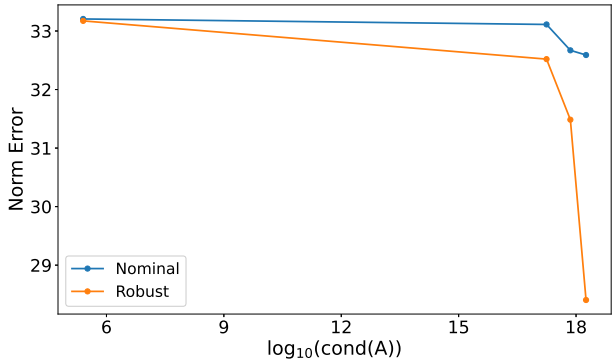
4.3. Numerical Experiments

In this section, we compare the robust eigen-pairs computed by Algorithm 2 with the nominal ones, while focusing on nearly singular matrices. We generate a matrix $\boldsymbol{A}_0 \in \mathbb{R}^{(n-1)\times n}$, with entries from $\mathcal{N}(0,1)$ and then form the matrix $\boldsymbol{A} = \boldsymbol{A}_0^T\boldsymbol{A}_0$. The metric used for evaluation is the average norm error on perturbed matrices. Namely, for a matrix $\boldsymbol{A}$, perturbed matrix $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{\Delta A}$ and vectors $\boldsymbol{\lambda}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ we compute $\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{B}\boldsymbol{x}_i - \lambda_i\boldsymbol{x}_i\|_2$. We take $\Delta A_{ij} \sim \lambda[0,1]$, $\lambda \in [1,5]$. The results are illustrated in Table 10.

Table 10: Columns Nominal, Rob.- $\boldsymbol{\ell_1}$ and Rob.- $\boldsymbol{\ell_2}$ correspond to the average norm error on matrices $\boldsymbol{A}+\boldsymbol{\Delta A}$ for nominal and robust, averaged over 100 matrices. Column $\|\boldsymbol{\Delta A}\|_{\boldsymbol{F}}$ refers to the average norm of $\boldsymbol{\Delta A}$ and the p-value corresponds to the null hypothesis of equal average norm errors for nominal and robust eigen-pairs.

$\boldsymbol{n}$	$\|\boldsymbol{\Delta A}\|_{\boldsymbol{F}}$	**Nominal**	**Rob.- $\boldsymbol{\ell_1}$**	**Rob.- $\boldsymbol{\ell_2}$**	**p-value**
10	43.41	12.48	**12.22**	12.37	10^{-3}
10	115.02	33.15	**32.41**	32.72	10^{-4}
10^2	433.55	39.25	39.02	**38.71**	10^{-4}
10^2	959.12	88.39	86.94	**86.82**	$< 10^{-5}$
3×10^2	2005.01	131.77	130.85	**129.55**	$< 10^{-5}$

Figure 4: The effect of cond($\boldsymbol{A}$) on the error of the nominal and robust eigen-pairs on $\boldsymbol{A} + \boldsymbol{\Delta A}$, averaged over 100 perturbed matrices. We vary cond($\boldsymbol{A}$), where $\boldsymbol{A} = \boldsymbol{A}_0^T \boldsymbol{A}_0$, by taking $\boldsymbol{A}_0 \in \mathbb{R}^{(n+k)\times n}$, for $k \in [-8, 10]$.



From Table 10, we observe that the robust eigen-pairs computed with Algorithm 2 are more accurate than the nominal on perturbations of ill conditioned matrices. Algorithm 2 converges in 10 iterations for $n = 10$ and in 30 iterations for $n = 100$. Moreover, it scales up to matrices with sizes in the hundreds, more precisely $n = 300$.

Further, from Figure 4 we notice that for large condition numbers the robust eigen-pairs are significantly more accurate than the nominal. This aligns with Theorem 6, since in such cases the matrix has a significant percentage of nearly zero eigenvalues.

Moreover, from Table 11 we observe that in all cases, the nominal approach requires on average less than a second. On the other hand, for $n = 10$ the robust approach requires on average 2.02 seconds with the ℓ_2 norm and 2.11

seconds with the ℓ_1 norm. Further, for $n = 100$ and $n = 300$, the robust approach with the ℓ_2 norm requires on average significantly less time that that with the ℓ_1 norm.

Table 11: Computational time comparison for nominal and robust eigen-pairs. All reported times are in seconds, averaged over 100 runs.

$\boldsymbol{n}$	**Nominal**	**Rob.- $\boldsymbol{\ell_1}$**	**Rob.- $\boldsymbol{\ell_2}$**
10	0.02 ± 0.01	2.11 ± 0.24	2.02 ± 0.11
10^2	0.03 ± 0.01	518.13 ± 51.84	176.31 ± 12.95
3×10^2	0.05 ± 0.03	4519.43 ± 89.07	3411.43 ± 17.75

5. Robust Matrix Factorizations

In this section, we derive an approach for computing matrix factorization under uncertainty. Based on this approach we compute the robust Cholesky factorization and compare it with the nominal.

5.1. Formulation

Finding an approximation for a given matrix, i.e. matrix completion, can be done by minimizing the norm error. For a nominal matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, it can be formulated as follows (Bertsimas and Copenhaver, 2018):

$$\min_{\boldsymbol{X}\in\mathcal{S}} \|\boldsymbol{A} - \boldsymbol{X}\|_F, \tag{18}$$

where $\mathcal{S} \subseteq \mathbb{R}^{n\times n}$ and contains problem specific constraints. We note that as (Bertsimas and Copenhaver, 2018) explained, applying the methodology introduced so far, that is,

$$\min_{\boldsymbol{X}\in\mathcal{S}} \max_{\boldsymbol{\Delta A}\in\mathcal{U}} \|\boldsymbol{A} + \boldsymbol{\Delta A} - \boldsymbol{X}\|_F,$$

where $\mathcal{U} \in \{\mathcal{U}_2, \mathcal{U}_1, \mathcal{U}_{11}\}$, is equivalent to simply changing the loss function. More precisely, if $g(\boldsymbol{Y}) = \|\boldsymbol{Y}\|_F$, then instead of having $\min_{\boldsymbol{X}} g(\boldsymbol{A} - \boldsymbol{X})$, we obtain $\min_{\boldsymbol{X}} \overline{g}(\boldsymbol{A} - \boldsymbol{X})$, where $\overline{g}(\boldsymbol{A} - \boldsymbol{X}) = \max_{\boldsymbol{\Delta A}\in\mathcal{U}} g((\boldsymbol{A} + \boldsymbol{\Delta A}) - \boldsymbol{X})$. Since g is a convex loss function, $\overline{g}$ is a new convex loss function of $\boldsymbol{A} - \boldsymbol{X}$.

Therefore, we do not treat this case here and instead introduce a different model of uncertainty. Let $\boldsymbol{\Delta}(\boldsymbol{X})$ be a linear mapping from $\mathbb{R}^{n\times n}$ to $\mathbb{R}^{n\times n}$,

defined by the matrices $\boldsymbol{\Delta}^{\boldsymbol{ij}}$, such that $\boldsymbol{\Delta}(\boldsymbol{X})_{ij} = \langle \boldsymbol{\Delta}^{\boldsymbol{ij}}, \boldsymbol{X} \rangle$. We then utilize the following spectral uncertainty set, see (Bertsimas and Copenhaver, 2018):

$$\mathcal{U}_{\sigma_p} = \{\boldsymbol{\Delta} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n} | \boldsymbol{\Delta} \text{ linear}, \ \|\boldsymbol{\Delta}\|_{\sigma_p} \leq \rho\},$$

where $\|\boldsymbol{\Delta}\|_{\sigma_p} = \max_{\boldsymbol{X}} \frac{\|\boldsymbol{\Delta}(\boldsymbol{X})\|_{\sigma_p}}{\|\boldsymbol{X}\|_{\sigma_p}}$. In this model of uncertainty we assume that $A_{ij} = X_{ij} + \sum_{kl} \Delta^{ij}_{kl} X_{kl} + \epsilon_{ij}$, where ϵ_{ij} denotes noise. We note that in this way we can take into account interactions among different entries of $\boldsymbol{X}$, see (Bertsimas and Copenhaver, 2018). Following the min-max robust approach, we propose minimizing the norm error under the worst possible realization of uncertainty within $\mathcal{U}_{\sigma_p}$, which for $p = 2$ is formulated as follows:

$$\min_{\boldsymbol{X}\in\mathcal{S}} \max_{\boldsymbol{\Delta}\in\mathcal{U}_{\sigma_2}} \|\boldsymbol{A} - \boldsymbol{X} - \boldsymbol{\Delta}(\boldsymbol{X})\|_F. \tag{19}$$

In this case, we have the following RC:

Theorem 9 ((Bertsimas and Copenhaver, 2018)). *The RC of Problem* (19) *is as follows:*

$$\min_{\boldsymbol{X}\in\mathcal{S}} \|\boldsymbol{A} - \boldsymbol{X}\|_F + \rho\|\boldsymbol{X}\|_F. \tag{20}$$

Utilizing the above methodology, we next show how we can compute a robust Cholesky factorization.

5.2. Robust Cholesky

Cholesky factorization is widely used in linear algebra and it consists of decomposing a positive definite matrix $\boldsymbol{A}$ as $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^T$ where $\boldsymbol{L}$ is lower triangular. This can be formulated as minimizing the norm error $\|\boldsymbol{A} - \boldsymbol{L}\boldsymbol{L}^T\|_F$, over the set of lower triangular matrices. More precisely, we define $\mathcal{S} = \{\boldsymbol{L} \in \mathbb{R}^{n\times n} : L_{ij} = 0, \ \ j > i\}$ and obtain the following problem:

$$\min_{\boldsymbol{L}\in\mathcal{S}} \|\boldsymbol{A} - \boldsymbol{L}\boldsymbol{L}^T\|_F \tag{21}$$

Utilizing the methodology introduced in Section 5.1, we obtain the following robust problem:

$$\min_{\boldsymbol{L}\in\mathcal{S}} \max_{\boldsymbol{\Delta}\in\mathcal{U}_{\sigma_2}} \|\boldsymbol{A} - \boldsymbol{L}\boldsymbol{L}^T - \boldsymbol{\Delta}(\boldsymbol{L}\boldsymbol{L}^T)\|_F, \tag{22}$$

for which the RC is as follows (see Theorem 9):

$$\min_{\boldsymbol{L}\in\mathcal{S}} \|\boldsymbol{A} - \boldsymbol{L}\boldsymbol{L}^T\|_F + \rho\|\boldsymbol{L}\boldsymbol{L}^T\|_F. \tag{23}$$

Problem (23) involves minimizing a non-convex function subject to linear constraints. Thus, we propose decomposing it into two separate optimization problems. First, we solve Problem (23) without the equality constraints, that is,

$$\min_{\boldsymbol{L}} \ \|\boldsymbol{A} - \boldsymbol{L}\boldsymbol{L}^T\|_F + \rho\|\boldsymbol{L}\boldsymbol{L}^T\|_F. \tag{24}$$

Let $\boldsymbol{L}^*$ denote the optimal solution of Problem (24). Then, we project $\boldsymbol{L}^*$ in $\mathcal{S}$ by solving the following problem:

$$\min_{\boldsymbol{L}\in\mathcal{S}} \|\boldsymbol{L}^* - \boldsymbol{L}\|_F^2. \tag{25}$$

A high quality solution $\boldsymbol{L}^*$ for the initial non-convex problem can be obtained with gradient based methods, i.e., the Adam algorithm (Kingma and Ba, 2015), initialized with the nominal Cholesky. Moreover, Problem (25) is a quadratic optimization problem, which can be solved efficiently with multiple solvers, i.e., Gurobi (Gurobi Optimization, LLC, 2021), OSQP (Stellato et al., 2020). We further note that the algorithm used for computing the nominal Cholesky factorization is stable and there is no dependence on the condition number of the matrix (Datta, 2010). The main benefit of the robust Cholesky factorization in this case is robustness, that is achieving smaller norm error on perturbed matrices than the nominal.

5.3. Numerical Experiments

In this section, we compare the robust with the nominal Cholesky factorization in terms of norm error on perturbed matrices. We generate a matrix $\boldsymbol{A}_0 \in \mathbb{R}^{n\times n}$ with entries from $[0, 1]$ and then form the matrix $\boldsymbol{A} = \boldsymbol{A}_0^T\boldsymbol{A}_0$. Then, we compute the nominal Cholesky factorization $\boldsymbol{L}_{\text{nom}}$ and the robust Cholesky factorization $\boldsymbol{L}_{\text{rob}}$. Problem (24) is solved with the Adam algorithm using 1000 iterations with learning rate 0.001 and Problem (25) is solved with Gurobi. Then, we generate perturbed matrices $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{\Delta A}$ and compute the norm errors $\|\boldsymbol{B} - \boldsymbol{L}_{\text{nom}}\boldsymbol{L}_{\text{nom}}^T\|_F$, $\|\boldsymbol{B} - \boldsymbol{L}_{\text{rob}}\boldsymbol{L}_{\text{rob}}^T\|_F$. We take $\Delta A_{ij} \sim \lambda[0, 1]$, $\lambda \in [1, 5]$, $\rho = 0.1$. The results are illustrated in Table 12.
As Table 12 illustrates, the robust Cholesky factorization is more accurate than the nominal, on perturbed matrices, across all dimensions. Moreover, we notice that the p-value is always less than 10^{-5} indicating that the difference in norm error between the nominal and robust Cholesky factorizations is statistically significant. Finally, we observe that in perturbed matrices the robust Cholesky factorization is more accurate than the nominal, independently of the matrix condition number.

Table 12: Nominal and robust Cholesky factorization norm error comparison averaged over 100 matrices and further for each one over 100 perturbed matrices. The p-value corresponds to the null hypothesis of equal average norm errors for nominal and robust Cholesky factorizations.

n	cond(A)	Nominal	Robust	p-value
10^2	10^3	172.49	**146.67**	$< 10^{-5}$
10^2	10^9	173.21	**145.44**	$< 10^{-5}$
10^2	10^{18}	173.01	**145.13**	$< 10^{-5}$
10^3	10^6	8647.64	**6330.93**	$< 10^{-5}$
10^3	10^{10}	8644.72	**6304.59**	$< 10^{-5}$
10^3	10^{18}	8706.74	**6339.86**	$< 10^{-5}$

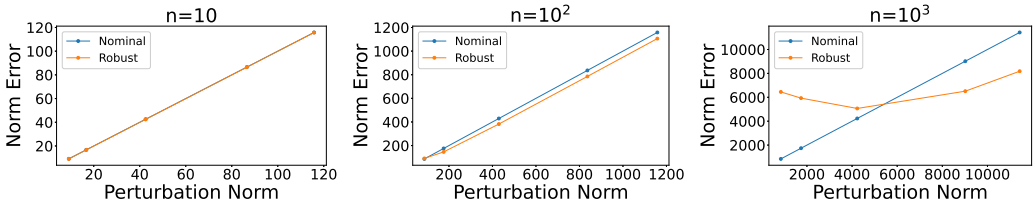


Figure 5: Comparison of nominal ($\boldsymbol{L}_{\text{nom}}$) and robust ($\boldsymbol{L}_{\text{rob}}$) Cholesky factorizations in terms of norm error on matrices $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{\Delta A}$. The x-axis represents the norm of the matrix $\boldsymbol{\Delta A}$ and the y-axis the norm errors, $\|\boldsymbol{B} - \boldsymbol{L}_{\text{nom}}\boldsymbol{L}_{\text{nom}}^T\|_F$ and $\|\boldsymbol{B} - \boldsymbol{L}_{\text{rob}}\boldsymbol{L}_{\text{rob}}^T\|_F$.

From Figure 5 we observe that as the dimension of the matrix and the amount of perturbation both increase, the robust Cholesky is more accurate than the nominal. More precisely, for $n = 10$ both the nominal and robust Cholesky achieve similar norm error on perturbed matrices. Further, for $n = 10^2$ they achieve similar norm error on small perturbations, while for large perturbations the robust Cholesky is more accurate, while for $n = 10^3$ the robust Cholesky is less accurate than the nominal on small perturbations and more accurate on large perturbations. Further, from Table 13 we observe that in all cases the robust Cholesky requires more computational time than the nominal, with the difference being more pronounced for $n = 10^3$.

6. Conclusion

In this paper, we addressed some of the fundamental problems in linear algebra, under the presence of uncertainty in the input data. We formulated

Table 13: Computational time comparison for nominal and robust Cholesky factorizations. All reported times are in seconds, averaged over 100 runs.

n	Nominal	Robust
10^2	$10^{-3} \pm 10^{-4}$	4.11 ± 0.13
10^3	0.01 ± 0.01	630.94 ± 15.35

them as computationally tractable optimization problems, illustrated properties that the robust solutions satisfy and also demonstrated their advantages over state of the art approaches in the numerical experiments. In a future work, it would be interesting to apply the derived framework for matrix factorization in other factorizations such as the QR factorization and also address other linear algebra problems in the presence of data uncertainty. Overall, we hope that the robust methods developed in this paper can be helpful in real world applications involving linear systems, matrix inversion, eigenvalues or factorizations under uncertainty in the input matrix.

Acknowledgements

We would like to thank the anonymous reviewers for providing us with insightful comments that helped improve the quality of the paper.

Appendix

We utilize the following result:

Theorem 10 ((Oettli and Prager, 1964)). *For any vector $\boldsymbol{x} \in \mathbb{R}^n$, there exists a matrix $\hat{\boldsymbol{A}} \in [\boldsymbol{A} - \boldsymbol{\Sigma}, \boldsymbol{A} + \boldsymbol{\Sigma}]$ such that $\boldsymbol{x}$ is the exact solution of the linear system $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$ if and only if $\boldsymbol{x}$ satisfies the following:*

$$|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}| \leq \boldsymbol{\Sigma}|\boldsymbol{x}|.$$

Proof of Theorem 2

Proof. Let $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$ for some $\boldsymbol{\Delta A} \in \mathcal{U}_{11}$. We have the following

$$|(\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b})_i| \leq \|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b}\|_1 \leq \bar{c}.$$

For the matrix $\boldsymbol{\Sigma}$, with i-th row defined as $\boldsymbol{\sigma}_i = \overline{c}\frac{|\overline{\boldsymbol{x}}|}{\|\overline{\boldsymbol{x}}\|_2^2}$, we observe that $\boldsymbol{\sigma}_i^T|\overline{\boldsymbol{x}}| = \overline{c}$, $\forall i$. Therefore, from Theorem 10 there exists a matrix $\hat{\boldsymbol{A}} \in \left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$ such that $\overline{\boldsymbol{x}}$ is the exact solution of the linear system $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$. Moreover, we have the following

$$\|\boldsymbol{\Delta A}\|_F^2 = \sum_{i=1}^{n} \|\boldsymbol{\Delta A}_i\|_2^2 \leq \sum_{i=1}^{n} \|\boldsymbol{\Delta A}_i\|_1^2 \leq \sum_{i=1}^{n} \rho^2 = n\rho^2.$$

Therefore, we obtain

$$\begin{aligned}
\|\hat{\boldsymbol{A}} - \boldsymbol{A}\|_F &= \|\hat{\boldsymbol{A}} - \overline{\boldsymbol{A}} + \overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \\
&\leq \|\hat{\boldsymbol{A}} - \overline{\boldsymbol{A}}\|_F + \|\overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \\
&\leq \|\boldsymbol{\Sigma}\|_F + \|\boldsymbol{\Delta A}\|_F \\
&= \sqrt{n}\frac{\overline{c}}{\|\overline{\boldsymbol{x}}\|_2} + \sqrt{n}\rho. \quad \square
\end{aligned}$$

We note that a similar result holds if $\overline{\boldsymbol{x}}$ is the optimal solution of problem (3) with optimal cost $\overline{c}$, since

$$|(\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b})_i| \leq \|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b}\|_1 \leq \sqrt{n}\|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}} - \boldsymbol{b}\|_2 \leq \sqrt{n}\overline{c}.$$

If we take the matrix $\boldsymbol{\Sigma}$ with i-th row $\boldsymbol{\sigma}_i = \sqrt{n}\overline{c}\frac{|\overline{\boldsymbol{x}}|}{\|\overline{\boldsymbol{x}}\|_2^2}$, we obtain $\boldsymbol{\sigma}_i^T|\overline{\boldsymbol{x}}| = \sqrt{n}\overline{c}$. Observe that the interval in this case is larger because of the $\sqrt{n}$ term.

Proof of Theorem 3

Proof. We observe that the family of matrices

$$\boldsymbol{B}(\theta) \;=\; (1 + \frac{\theta}{\|\boldsymbol{A}\|_F})\boldsymbol{A}, \;\; \theta \in [0, \rho],$$

are of the form $\boldsymbol{A} + \boldsymbol{\Delta A}$, $\boldsymbol{\Delta A} \in \mathcal{U}_2$. Therefore, we have the following:

$$\begin{aligned}
\|\boldsymbol{M}(\theta)\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{M}(\theta)\boldsymbol{b}\|_2 &= \|\boldsymbol{B}(\theta)\overline{\boldsymbol{x}} - (1 + \frac{\theta}{\|\boldsymbol{A}\|_F})\boldsymbol{b}\|_2 \\
&= \|\boldsymbol{B}(\theta)\overline{\boldsymbol{x}} - \boldsymbol{b} - \frac{\theta}{\|\boldsymbol{A}\|_F}\boldsymbol{b}\|_2 \\
&\leq \|\boldsymbol{B}(\theta)\overline{\boldsymbol{x}} - \boldsymbol{b}\|_2 + \|\frac{\theta}{\|\boldsymbol{A}\|_F}\boldsymbol{b}\|_2 \\
&\leq \overline{c} + \theta\frac{\|\boldsymbol{b}\|_2}{\|\boldsymbol{A}\|_F}. \quad \square
\end{aligned}$$

Proof of Theorem 8

PROOF. Let $\overline{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\Delta A}$ for some $\boldsymbol{\Delta A} \in \mathcal{U}_2$. We have the following

$$|(\overline{\boldsymbol{A}}\overline{\boldsymbol{x}}_i - \overline{\lambda}_i\overline{\boldsymbol{x}}_i)_j| \leq \|\overline{\boldsymbol{A}}\overline{\boldsymbol{x}}_i - \overline{\lambda}_i\overline{\boldsymbol{x}}_i\|_1 \leq \overline{c}_i.$$

For the matrix $\boldsymbol{\Sigma}$, with j-th row defined as $\boldsymbol{\sigma}_j = \overline{c}_i|\overline{\boldsymbol{x}}_i|$, we observe that $\boldsymbol{\sigma}_j^T|\overline{\boldsymbol{x}}_i| = \overline{c}_i$, $\forall j$. Therefore, from Theorem 10 there exists a matrix $\hat{\boldsymbol{A}} \in \left[\overline{\boldsymbol{A}} - \boldsymbol{\Sigma}, \overline{\boldsymbol{A}} + \boldsymbol{\Sigma}\right]$ such that $\hat{\boldsymbol{A}}\overline{\boldsymbol{x}}_i = \overline{\lambda}_i\overline{\boldsymbol{x}}_i$. Moreover, we obtain

$$\begin{aligned}
\|\hat{\boldsymbol{A}} - \boldsymbol{A}\|_F &= \|\hat{\boldsymbol{A}} - \overline{\boldsymbol{A}} + \overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \\
&\leq \|\hat{\boldsymbol{A}} - \overline{\boldsymbol{A}}\|_F + \|\overline{\boldsymbol{A}} - \boldsymbol{A}\|_F \\
&\leq \|\boldsymbol{\Sigma}\|_F + \|\boldsymbol{\Delta A}\|_F \\
&= \sqrt{n}\overline{c}_i + \sqrt{n}\rho. \quad \square
\end{aligned}$$

References

Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. Robust Optimization. volume 28. Princeton University Press.

Bertsimas, D., Copenhaver, M.S., 2018. Characterization of the equivalence of robustification and regularization in linear and matrix regression. European Journal of Operational Research 270, 931–942.

Bertsimas, D., den Hertog, D., 2022. Robust and Adaptive Optimization. Dynamic Ideas.

Bertsimas, D., Ma, Y., 2023. Robust Regression over Averaged Uncertainty. Working Paper .

Björck, Å., 1991. Component-wise perturbation analysis and error bounds for linear least squares solutions. BIT Numerical Mathematics 31, 237–244.

Candès, E.J., Li, X., Ma, Y., Wright, J., 2011. Robust principal component analysis? Journal of the ACM (JACM) 58, 1–37.

Choi, S.C., 2006. Iterative methods for singular linear equations and least-squares problems. Ph.D. thesis.

Corsaro, S., Marino, M., 2006. Interval linear systems: the state of the art. Computational Statistics 21, 365–384.

Daniel, J.W., 1967. The conjugate gradient method for linear and nonlinear operator equations. SIAM Journal on Numerical Analysis 4, 10–26.

Datta, B.N., 2010. Numerical linear algebra and applications. volume 116. SIAM.

Doyle, J., Newlin, M., Paganini, F., Tierno, J., 1994. Unifying robustness analysis and system id, in: Proceedings of 1994 33rd IEEE Conference on Decision and Control, IEEE. pp. 3667–3672.

El Ghaoui, L., 2002. Inversion error, condition number, and approximate inverses of uncertain matrices. Linear algebra and its applications 343, 171–193.

El Ghaoui, L., Lebret, H., 1997. Robust solutions to least-squares problems with uncertain data. SIAM Journal on Matrix Analysis and Applications 18, 1035–1064.

Eldén, L., 1980. Perturbation theory for the least squares problem with linear equality constraints. SIAM Journal on Numerical Analysis 17, 338–350.

Fiedler, M., Nedoma, J., Ramík, J., Rohn, J., Zimmermann, K., 2006. Linear optimization problems with inexact data. Springer Science & Business Media.

Fierro, R.D., Bunch, J.R., 1994. Collinearity and total least squares. SIAM Journal on Matrix Analysis and Applications 15, 1167–1181.

Fong, D.C.L., Saunders, M., 2011. Lsmr: An iterative algorithm for sparse least-squares problems. SIAM Journal on Scientific Computing 33, 2950–2971.

Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9, 432–441.

Friedman, J., Hastie, T., Tibshirani, R., Tibshirani, M.R., 2015. Package 'glasso'.

Golub, G.H., Van Loan, C.F., 1980. An analysis of the total least squares problem. SIAM Journal on Numerical Analysis 17, 883–893.

Gurobi Optimization, LLC, 2021. Gurobi Optimizer Reference Manual. URL: https://www.gurobi.com.

Hager, W.W., 1989. Updating the inverse of a matrix. SIAM review 31, 221–239.

Higham, N.J., 2002. Accuracy and stability of numerical algorithms. SIAM.

Hong, M., Wang, X., Razaviyayn, M., Luo, Z.Q., 2017. Iteration complexity analysis of block coordinate descent methods. Mathematical Programming 163, 85–114.

Huang, J., Nie, F., Huang, H., Ding, C., 2014. Robust manifold nonnegative matrix factorization. ACM Transactions on Knowledge Discovery from Data (TKDD) 8, 1–21.

Kingma, D.P., Ba, J., 2015. Adam: A method for stochastic optimization, in: 3rd International Conference on Learning Representations, ICLR.

Knyazev, A.V., 2001. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. SIAM Journal on Scientific Computing 23, 517–541.

Kreinovich, V., Lakeyev, A.V., Rohn, J., Kahl, P., 2013. Computational complexity and feasibility of data processing and interval computations. volume 10. Springer Science & Business Media.

Netrapalli, P., UN, N., Sanghavi, S., Anandkumar, A., Jain, P., 2014. Non-convex robust pca. Advances in neural information processing systems 27.

Oettli, W., Prager, W., 1964. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. Numerische Mathematik 6, 405–409.

Ogita, T., Oishi, S., 2012. Accurate and robust inverse cholesky factorization. Nonlinear Theory and Its Applications, IEICE 3, 103–111.

Paige, C.C., Saunders, M.A., 1975. Solution of sparse indefinite systems of linear equations. SIAM Journal on Numerical Analysis 12, 617–629.

Saad, Y., Schultz, M.H., 1986. Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM Journal on Scientific and Statistical Computing 7, 856–869.

Scherzinger, W.M., Dohrmann, C.R., 2008. A robust algorithm for finding the eigenvalues and eigenvectors of $3\times$ 3 symmetric matrices. Computer Methods in Applied Mechanics and Engineering 197, 4007–4015.

Stellato, B., Banjac, G., Goulart, P., Bemporad, A., Boyd, S., 2020. OSQP: An operator splitting solver for quadratic programs. Mathematical Programming Computation 12, 637–672.

Sun, R., Hong, M., 2015. Improved iteration complexity bounds of cyclic block coordinate descent for convex problems. Advances in Neural Information Processing Systems 28.

Trefethen, L.N., Bau III, D., 1997. Numerical linear algebra. volume 50. SIAM.

Weyl, H., 1912. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). Mathematische Annalen 71, 441–479.

Xu, H., Caramanis, C., Mannor, S., 2010. Robust regression and Lasso. IEEE Transactions on Information Theory 56, 3561–3574.

Yuan, M., Lin, Y., 2007. Model selection and estimation in the gaussian graphical model. Biometrika 94, 19–35.

Zhen, J., den Hertog, D., 2017. Centered solutions for uncertain linear equations. Computational Management Science 14, 585–610.