# Convex Optimization Problems

Cong Ma

University of Chicago, Winter 2026

# Outline

- Convex optimization: two equivalent forms
- Equivalent convex problems (reformulations)
- Linear programming (LP) and an application (compressed sensing)
- Quadratic programming (QP) and an application (SVM)
- Quadratically constrained QP (QCQP) and hidden convexity (TRS)
- Semidefinite programming (SDP) and applications (eigenvalue / norms / NNM)

# Convex Optimization: General Form

A convex optimization problem (or convex problem) has the form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \mathcal{C}, \end{aligned}$$

where

- $\mathcal{C} \subseteq \mathbb{R}^n$ is a **convex set**,
- $f : \mathcal{C} \to \mathbb{R}$ is a **convex function**.

# Convex Optimization: Functional Form

A common representation is

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0, \quad i = 1, \ldots, m, \\ & h_j(\boldsymbol{x}) = 0, \quad j = 1, \ldots, p, \end{aligned}$$

where $f, g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ are **convex**, and $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are **affine**.

**Key fact**

The feasible set $\{\boldsymbol{x} : g_i(\boldsymbol{x}) \le 0,\ h_j(\boldsymbol{x}) = 0\}$ is convex.

# Key Property of Convex Problems

**Theorem 1**

*Let $\mathcal{C} \subseteq \mathbb{R}^n$ be convex and $f : \mathcal{C} \to \mathbb{R}$ be convex. If $\boldsymbol{x}^\star \in \mathcal{C}$ is a local minimum of $f$ over $\mathcal{C}$, then $\boldsymbol{x}^\star$ is a global minimum.*

# Equivalent Convex Problems

Two problems are (informally) **equivalent** if the solution of one can be readily obtained from the solution of the other.

**Common convexity-preserving transformations:**

- eliminate affine equality constraints,
- introduce slack variables,
- epigraph reformulations,
- introduce auxiliary variables for affine compositions,
- partial minimization.

# Equivalent Convex Problems: Variable Transformations

**1. Eliminating Equality Constraints** Minimize $f_0(\boldsymbol{x})$ s.t. $f_i(\boldsymbol{x}) \leq 0$ and $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is equivalent to:

$$\min_{\boldsymbol{z}} f_0(\boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0) \quad \text{s.t.} \quad f_i(\boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0) \leq 0$$

where $\boldsymbol{x} = \boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0$ parameterizes the affine set $\{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$.

**2. Introducing Slack Variables** Replacing linear inequalities $\boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i$ with equalities:

$$\min_{\boldsymbol{x},\boldsymbol{s}} f_0(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{a}_i^\top \boldsymbol{x} + s_i = b_i, \quad s_i \geq 0$$

# Equivalent Convex Problems: Functional Forms

**3. Epigraph Form** Standard form is equivalent to minimizing a linear objective over the epigraph:

$$\min_{\boldsymbol{x},t} \quad t \quad \text{s.t.} \quad f_0(\boldsymbol{x}) - t \le 0, \quad f_i(\boldsymbol{x}) \le 0, \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

**4. Introducing Equality Constraints** Transforming $f_i(\boldsymbol{A}_i\boldsymbol{x} + \boldsymbol{b}_i)$ by letting $\boldsymbol{y}_i = \boldsymbol{A}_i\boldsymbol{x} + \boldsymbol{b}_i$:

$$\min_{\boldsymbol{x},\boldsymbol{y}_i} f_0(\boldsymbol{y}_0) \quad \text{s.t.} \quad f_i(\boldsymbol{y}_i) \le 0, \quad \boldsymbol{y}_i = \boldsymbol{A}_i\boldsymbol{x} + \boldsymbol{b}_i$$

**5. Partial Minimization** If $f_0(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is convex in $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ and $C$ is a convex set:

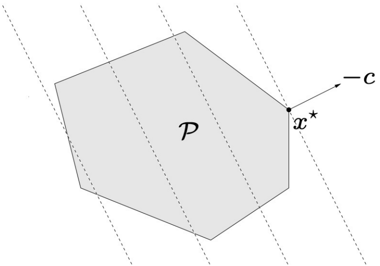
$$\min f_0(\boldsymbol{x}_1, \boldsymbol{x}_2) \text{ s.t. } \boldsymbol{x}_1 \in C \iff \min \tilde{f}_0(\boldsymbol{x}_1) \text{ s.t. } \boldsymbol{x}_1 \in C$$

where $\tilde{f}_0(\boldsymbol{x}_1) = \inf_{\boldsymbol{x}_2} f_0(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the partial minimization.

# Linear Program (LP)

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + d \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{x} \le \boldsymbol{h}, \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{aligned}$$

- objective and constraint functions are **affine**,
- feasible set is a **polyhedron**.



# Compressed Sensing and Linear Programming

**Problem Setting.** Recover a sparse signal $\boldsymbol{x} \in \mathbb{R}^n$ from a small number of linear measurements $\boldsymbol{y} \in \mathbb{R}^m$ with $m \ll n$:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

**Key Assumptions:**

- **Sparsity:** $\boldsymbol{x}$ has at most $k$ nonzero entries, with $k \ll n$.
- **Incoherence / RIP:** The measurement matrix $\boldsymbol{A}$ satisfies suitable conditions (e.g., Restricted Isometry Property).

**Convex Relaxation (Basis Pursuit).** Direct $\ell_0$-minimization is NP-hard. Instead, solve the convex problem

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\boldsymbol{x}\|_1 \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \end{aligned}$$

# Reformulation as a Linear Program

The $\ell_1$ minimization problem can be rewritten as a **linear program** by introducing auxiliary variables $\boldsymbol{u} \in \mathbb{R}^n$:

**Equivalent LP Form:**

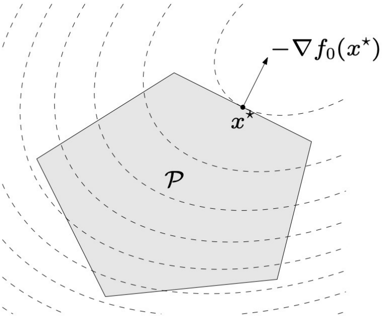
$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{u}} \quad & \sum_{i=1}^{n} u_i \\
\text{s.t.} \quad & -u_i \le x_i \le u_i, \quad i = 1, \ldots, n, \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}.
\end{aligned}$$

- The objective and constraints are **affine** $\Rightarrow$ this is a convex LP.
- Any local optimum is therefore a **global optimum**.
- Enables the use of highly efficient LP solvers for large-scale signal recovery.

# Quadratic Program (QP)

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{x} \le \boldsymbol{h}, \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{aligned}$$

- $\boldsymbol{P} \in \mathbb{S}^n_+$ so the objective is **convex quadratic**,
- minimizing a convex quadratic over a polyhedron.



# Linear Classifier and Margins

Given labeled data $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^m$ with $y_i \in \{-1, 1\}$, a linear classifier uses a hyperplane

$$\boldsymbol{w}^\top \boldsymbol{x} + b = 0.$$

**Functional margin of a single example:**

$$\hat{\gamma}^{(i)} = y_i\big(\boldsymbol{w}^\top \boldsymbol{x}_i + b\big).$$

**Geometric margin:** distance of $(\boldsymbol{x}_i, y_i)$ to the hyperplane is

$$\gamma^{(i)} = \frac{y_i\big(\boldsymbol{w}^\top \boldsymbol{x}_i + b\big)}{\|\boldsymbol{w}\|_2}.$$

This is invariant to scaling of $(\boldsymbol{w}, b)$ and thus the meaningful measure of confidence.

# Maximum Margin Optimization

The **minimum geometric margin over all training points** is

$$\gamma = \min_{i=1,\ldots,m} \gamma^{(i)} = \min_i \frac{y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)}{\|\boldsymbol{w}\|_2}.$$

To find the best separating hyperplane, we maximize this minimum margin:

$$\max_{\boldsymbol{w},b} \ \gamma \quad \text{s.t.} \quad y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq \gamma, \quad i = 1, \ldots, m.$$

However, this is not yet a standard convex program because $\|\boldsymbol{w}\|_2$ appears in the denominator and in the objective.

# Scaling and Canonical Constraints

Because geometric margin is invariant to scaling of $(\boldsymbol{w}, b)$, we can **fix the functional margin** to 1 at the support vectors:

$$y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1, \quad i = 1, \ldots, m.$$

Under this constraint, the geometric margin becomes

$$\gamma = \frac{1}{\|\boldsymbol{w}\|_2}.$$

Thus maximizing $\gamma$ is equivalent to minimizing $\|\boldsymbol{w}\|_2$, which we write as minimizing the squared norm for smoothness:

$$\min_{\boldsymbol{w}, b} \ \frac{1}{2}\|\boldsymbol{w}\|_2^2 \quad \text{s.t.} \quad y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1.$$

This is a classical formulation of the **hard-margin support vector machine**.

# SVM as a Quadratic Program

The hard-margin SVM optimization problem is

$$\begin{aligned}\min_{\boldsymbol{w},b} \quad & \tfrac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} \\ \text{s.t.} \quad & y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1, \quad i = 1, \ldots, m.\end{aligned}$$

**Why this is a QP:**

- The objective $\frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w}$ is a convex quadratic form.
- The constraints $y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1$ are **affine** in $(\boldsymbol{w}, b)$.

Thus the SVM problem fits the standard convex QP template:

$$\min_{\boldsymbol{x}} \ \tfrac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} \quad \text{s.t. affine constraints,}$$

with $\boldsymbol{x} = (\boldsymbol{w}, b)$.

# Quadratically Constrained Quadratic Program (QCQP)

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_0 \boldsymbol{x} + \boldsymbol{q}_0^\top \boldsymbol{x} + r_0 \\
\text{s.t.} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} + \boldsymbol{q}_i^\top \boldsymbol{x} + r_i \leq 0, \quad i = 1, \ldots, m, \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.
\end{aligned}
$$

- if $\boldsymbol{P}_i \in \mathbb{S}_{+}^{n}$, then objective/constraints are **convex quadratics**,
- if $\boldsymbol{P}_i \in \mathbb{S}_{++}^{n}$, feasible region is an intersection of ellipsoids and an affine set.

# Hidden Convexity in Trust Region Subproblems

## Problem Formulation (TRS)

$$\min\{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + 2\boldsymbol{b}^\top \boldsymbol{x} + c : \|\boldsymbol{x}\|^2 \le 1\}$$

where $\boldsymbol{b} \in \mathbb{R}^n, c \in \mathbb{R}$ and $\boldsymbol{A} \in \mathbb{S}^n$. **Generally non-convex.**

## Spectral Decomposition

Let $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\top$ with orthogonal $\boldsymbol{U}$ and diagonal $\boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_n)$. Substituting into (TRS) and using $\|\boldsymbol{U}^\top \boldsymbol{x}\|^2 = \|\boldsymbol{x}\|^2$:

$$\min\{\boldsymbol{x}^\top \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\top \boldsymbol{x} + 2\boldsymbol{b}^\top \boldsymbol{U}\boldsymbol{U}^\top \boldsymbol{x} + c : \|\boldsymbol{U}^\top \boldsymbol{x}\|^2 \le 1\}$$

## Change of Variables

Set $\boldsymbol{y} = \boldsymbol{U}^\top \boldsymbol{x}$ and $\boldsymbol{f} = \boldsymbol{U}^\top \boldsymbol{b}$. The problem reduces to:

$$\begin{aligned} \min \quad & \sum_{i=1}^n d_i y_i^2 + 2\sum_{i=1}^n f_i y_i + c \\ \text{s.t.} \quad & \sum^n y_i^2 \le 1 \end{aligned} \tag{2}$$

# Hidden Convexity in Trust Region Subproblems Contd.

**Lemma 2**

*Let $\boldsymbol{y}^*$ be an optimal solution of (2). Then $f_i y_i^* \leq 0$ for all $i = 1, 2, \ldots, n$.*

**Proof.** Denote the objective function of (2) by $g(\boldsymbol{y}) \equiv \sum_{i=1}^n d_i y_i^2 + 2\sum_{i=1}^n f_i y_i + c$. Let $i \in \{1, 2, \ldots, n\}$. Define $\tilde{\boldsymbol{y}}$ as

$$\tilde{y}_j = \begin{cases} y_j^* & j \neq i, \\ -y_i^* & j = i. \end{cases}$$

$\tilde{\boldsymbol{y}}$ is feasible and $g(\boldsymbol{y}^*) \leq g(\tilde{\boldsymbol{y}})$.

$$\sum_{i=1}^n d_i (y_i^*)^2 + 2\sum_{i=1}^n f_i y_i^* + c \leq \sum_{i=1}^n d_i (\tilde{y}_i)^2 + 2\sum_{i=1}^n f_i \tilde{y}_i + c.$$

After cancellation of terms, $2f_i y_i^* \leq 2f_i(-y_i^*)$, implying the desired inequality $f_i y_i^* \leq 0$.

# Back to the TRS problem

Using the lemma $f_i y_i^* \leq 0$, we make the change of variable:

$$y_i = -\mathsf{sgn}(f_i)\sqrt{z_i} \quad (z_i \geq 0).$$

Problem (2) then becomes:

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} d_i z_i - 2\sum_{i=1}^{n} |f_i|\sqrt{z_i} + c \\ \text{s.t.} \quad & \sum_{i=1}^{n} z_i \leq 1, \\ & z_1, z_2, \ldots, z_n \geq 0. \end{aligned}$$

This is now a **convex optimization problem** because:

- The term $-\sqrt{z_i}$ is a convex function for $z_i \geq 0$.
- The constraints define a convex set (a simplex-like region).

# Semidefinite Program (SDP)

**Standard form:**

$$
\begin{aligned}
\min_{X} \quad & \mathbf{tr}(CX) \\
\text{s.t.} \quad & \mathbf{tr}(A_i X) = b_i, \quad i = 1, \dots, m, \qquad\qquad X, C, A_i \in \mathbb{S}^n \\
& X \succeq 0,
\end{aligned}
$$

- Optimization variable is a **matrix**.
- Linear objective + affine constraints.
- PSD constraint defines a **convex cone**.

**Geometry**

Linear function minimized over an affine slice of the PSD cone.

# LMI Formulation

Equivalent SDP form:

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.} \quad & F(x) \preceq 0, \\ & Ax = b, \end{aligned} \qquad F(x) = G + \sum_{i=1}^{n} x_i F_i$$

with $F_i, G \in \mathbb{S}^k$.

- $F(x) \preceq 0$ is a **linear matrix inequality (LMI)**.
- Decision variable is now a **vector**.
- Same convex geometry.

# Standard Form vs LMI Form

Standard SDP:

$$\min_X \ \mathbf{tr}(CX) \quad \text{s.t. } X \in \mathcal{F}, \ X \succeq 0$$

where

$$\mathcal{F} = \{X : \mathbf{tr}(A_i X) = b_i\}$$

Any $X \in \mathcal{F}$ can be written as

$$X = G + \sum_j x_j F_j$$

Substitute into constraint:

$$G + \sum_j x_j F_j \succeq 0$$

**Key Idea**

LMI form is a coordinate parameterization of the affine feasible set.

# Multiple LMI Constraints

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & x_1F_1 + x_2F_2 + \cdots + x_nF_n + G \preceq 0 \\ & Ax = b \end{array}$$

with $F_i, G \in \mathbb{S}^k$

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1\widehat{F}_1 + \cdots + x_n\widehat{F}_n + \widehat{G} \preceq 0, \quad x_1\widetilde{F}_1 + \cdots + x_n\widetilde{F}_n + \widetilde{G} \preceq 0$$

  is equivalent to single LMI

$$x_1\begin{bmatrix} \widehat{F}_1 & \\ & \widetilde{F}_1 \end{bmatrix} + x_2\begin{bmatrix} \widehat{F}_2 & \\ & \widetilde{F}_2 \end{bmatrix} + \cdots + x_n\begin{bmatrix} \widehat{F}_n & \\ & \widetilde{F}_n \end{bmatrix} + \begin{bmatrix} \widehat{G} & \\ & \widetilde{G} \end{bmatrix} \preceq 0$$

# LP as a Special Case of SDP (LMI View)

Consider the linear program

$$\min_x \; c^\top x \qquad \text{s.t. } Ax \le b$$

Each scalar constraint

$$a_i^\top x \le b_i$$

can be written as the $1 \times 1$ LMI:

$$\left[ a_i^\top x - b_i \right] \preceq 0$$

since $\mathbb{S}^1 = \mathbb{R}$.

Stacking all constraints gives a diagonal LMI:

$$\operatorname{diag}(Ax - b) \preceq 0$$

**Conclusion**

LP is an SDP with only **diagonal symmetric matrix variables**.

# Important SDP Modeling Examples

- Eigenvalue optimization
- Matrix norm minimization
- Nuclear norm minimization

# Eigenvalue minimization

$$\text{minimize} \quad \lambda_{\max}(A(x))$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbb{S}^k$)

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x) \preceq tI \end{array}$$

- variables $x \in \mathbb{R}^n$, $t \in \mathbb{R}$
- follows from

$$\lambda_{\max}(A) \leq t \qquad \Longleftrightarrow \qquad A \preceq tI$$

# Matrix norm minimization

$$\text{minimize} \quad \|A(x)\| = \left(\lambda_{\max}(A(x)^\top A(x))\right)^{1/2}$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbb{R}^{p\times q}$)

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} tI & A(x) \\ A(x)^\top & tI \end{array}\right] \succeq 0 \end{array}$$

- variables $x \in \mathbb{R}^n$, $t \in \mathbb{R}$
- constraint follows from

$$\begin{aligned} \|A\| \leq t \quad &\Longleftrightarrow \quad A^\top A \preceq t^2 I, \quad t \geq 0 \\ &\Longleftrightarrow \quad \left[\begin{array}{cc} tI & A \\ A^\top & tI \end{array}\right] \succeq 0 \end{aligned}$$

# Why Nuclear Norm Minimization?

Goal: recover a low-rank matrix

$$\min_X \ \mathrm{rank}(X) \quad \text{s.t. } \mathcal{A}(X) = y$$

Problems:

- Rank minimization is **nonconvex** and NP-hard
- Hard to optimize directly

Convex relaxation:

$$\min_X \ \|X\|_* \quad \text{s.t. } \mathcal{A}(X) = y$$

where

$$\|X\|_* = \sum_i \sigma_i(X)$$

# Why Nuclear Norm Minimization Is an SDP

Key identity:

$$\|X\|_* = \min_{W_1, W_2} \; \frac{1}{2}(\mathbf{tr}\, W_1 + \mathbf{tr}\, W_2) \quad \text{s.t.} \quad \begin{bmatrix} W_1 & X \\ X^\top & W_2 \end{bmatrix} \succeq 0$$

Therefore NNM

$$\min_X \|X\|_* \quad \text{s.t. } \mathcal{A}(X) = y$$

is equivalent to the SDP

$$\begin{aligned} \min_{X, W_1, W_2} \quad & \frac{1}{2}(\mathbf{tr}\, W_1 + \mathbf{tr}\, W_2) \\ \text{s.t.} \quad & \begin{bmatrix} W_1 & X \\ X^\top & W_2 \end{bmatrix} \succeq 0, \quad \mathcal{A}(X) = y \end{aligned}$$

# Nuclear Norm SDP Identity

**Claim:** For any $X \in \mathbb{R}^{m\times n}$,

$$\|X\|_* = \min_{W_1, W_2} \ \frac{1}{2}(\mathbf{tr}\, W_1 + \mathbf{tr}\, W_2)$$

subject to

$$\begin{bmatrix} W_1 & X \\ X^\top & W_2 \end{bmatrix} \succeq 0$$

We prove by showing:

- (Upper bound) Construct feasible $(W_1, W_2)$ achieving $\|X\|_*$
- (Lower bound) Any feasible solution has objective $\geq \|X\|_*$

# Proof (Upper Bound)

Let SVD of $X$ be

$$X = U\Sigma V^\top, \qquad \Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$$

Choose

$$W_1 = U\Sigma U^\top, \qquad W_2 = V\Sigma V^\top$$

Then

$$\begin{bmatrix} W_1 & X \\ X^\top & W_2 \end{bmatrix} = \begin{bmatrix} U\sqrt{\Sigma} \\ V\sqrt{\Sigma} \end{bmatrix} \begin{bmatrix} U\sqrt{\Sigma} \\ V\sqrt{\Sigma} \end{bmatrix}^\top \succeq 0$$

and

$$\frac{1}{2}(\mathbf{tr}\, W_1 + \mathbf{tr}\, W_2) = \mathbf{tr}(\Sigma) = \|X\|_*$$

So the SDP value is **at most** $\|X\|_*$.

# Proof (Lower Bound)

Let

$$\begin{bmatrix} W_1 & X \\ X^\top & W_2 \end{bmatrix} \succeq 0$$

Then by PSD block matrix property:

$$u^\top X v \leq \frac{1}{2}(u^\top W_1 u + v^\top W_2 v)$$

for all vectors $u, v$.

Take $u, v$ to be singular vector pairs of $X$ and sum over all singular values:

$$\sum_i \sigma_i(X) \leq \frac{1}{2}(\mathbf{tr}\, W_1 + \mathbf{tr}\, W_2)$$

Therefore every feasible point satisfies

$$\text{objective} \geq \|X\|_*$$