ECE285: Semidefinite and sum-of-squares optimization **Winter 2023**

Lecture 5: Conic programming

Lecturer: Yang Zheng *Scribe: Yang Zheng*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. Any typos should be sent to* `zhengy@eng.ucsd.edu`.

Learning goals:

1. General conic programming
2. Linear programming
3. Semidefinite programming

In this lecture, we introduce conic programs which are a far-reaching generalization of linear programs. Conic programs include a hierarchy of convex optimization problems: linear programs (LPs), quadratic programs (QPs), quadratically constrained quadratic programs (QCQPs), second-order cone programs (SOCPs), and semidefinite programs (SDPs), which will be discussed in Lecture 6. We present their standard forms and some selected examples.

1 General conic programming

Let $K \subseteq \mathbb{R}^n$ be a proper cone (closed, convex, non-empty interior, and pointed). A *standard conic program* has the form

$$\begin{aligned} \min \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Ax = b, \\ & x \in K, \end{aligned} \tag{1}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ are problem data, and the optimization variable is $x \in \mathbb{R}^n$. The feasible region is the set of $x \in \mathbb{R}^n$ satisfying the constraints $x \in K$ and $Ax = b$. In other words, the feasible set is the intersection of the *proper cone* K with the *affine subspace* $\{x \in \mathbb{R}^n \mid Ax = b\}$ and thus is a closed convex set.

An *inequality-form* conic program is an optimization problem of the form

$$\begin{aligned} \min \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Fx + g \preceq_K 0, \end{aligned} \tag{2}$$

where $F \in \mathbb{R}^{n\times m}, g \in \mathbb{R}^n, c \in \mathbb{R}^m$ are problem data, and the optimization variable is $x \in \mathbb{R}^m$ (the dimension here is different from that in (1)). The feasible set is the pre-image of the proper cone K under the affine function $f(x) = -Fx - g$, i.e., $f^{-1}(K) = \{x \mid f(x) = -Fx - g \in K\}$, which is thus a closed convex set (the closedness comes from the fact that the pre-image of a continuous function on a closed set is closed).

The standard-form conic program (1) and the inequality-form conic program (2) can be converted from each other equivalently. The conversion often involves adding or introducing new variables, leading to potentially different dimensions of the proper cone K (note that this conversion is very different from the duality that will be discussed in Lectures 7 & 8). We present some conversion details in Section 2 and Section 3.

2 Linear programming

2.1 Standard and inequality forms

A linear program (LP) is a conic program over the nonnegative orthant $K = \mathbb{R}^n_+$. Accordingly, the *standard-form LP* is a problem of the form

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{3}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ are problem data, and $x \geq 0$ denotes component-wise inequalities, $x_i \geq 0, i = 1, \ldots, n$. The feasible region of the LP (3) is defined by a set of finite linear inequalities and equations, which is called a *polyhedron*. Geometrically, an LP exactly corresponds to the minimization of a linear function over a polyhedron.

Example 5.1. *The following optimization problem is an LP*

$$\begin{aligned} \min_{x} \quad & x_1 + x_2 \\ \textit{subject to} \quad & 2x_1 - x_2 = 1, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

This problem is an instance of (3) *where* $x \in \mathbb{R}^2$, $K = \mathbb{R}^2_+$, *the cost vector is* $c = [1, 1]^\mathsf{T}$, *the matrix* A *and vector* b *are given by* $A = [2, -1]$ *and* $b = 1$. □

The *inequality-form LP* is an optimization problem written as

$$\begin{aligned} \min_{z\in\mathbb{R}^k} \quad & e^\mathsf{T} z \\ \text{subject to} \quad & Fz + g \geq 0, \end{aligned} \tag{4}$$

where $e \in \mathbb{R}^k, F \in \mathbb{R}^{n\times k}$ and $g \in \mathbb{R}^n$ are problem data, and the proper cone is $K = \mathbb{R}^n_+$. The standard-form LP (3) and the inequality-form LP (4) can be converted equivalently from each other:

- To go from (3) to (4), we eliminate the affine constraint in (3). In particular, choose a point g in the affine subspace $\{x \in \mathbb{R}^n \mid Ax = b\}$, and let $F \in \mathbb{R}^{n\times k}$ be the matrix whose range is the same as $\ker(A)$, where $k = \dim\ker(A)$. Then we have

  $$\{x \in \mathbb{R}^n \mid Ax = b\} = \{Fz + g \mid z \in \mathbb{R}^k\}.$$

  Thus, (3) is equivalent to

  $$\begin{aligned} \min_{z\in\mathbb{R}^k} \quad & c^\mathsf{T}(Fz + g) \\ \text{subject to} \quad & Fz + g \geq 0. \end{aligned} \tag{5}$$

  We see that (5) is equivalent to the inequality-form LP (4) by letting $e = F^\mathsf{T} c$ and noting that $c^\mathsf{T} g$ is a constant.

- To go from (4) to (3), we add slack variables s_i for the inequalities, leading to

  $$\begin{aligned} \min_{z,s} \quad & e^\mathsf{T} z \\ \text{subject to} \quad & Fz + g = s, \\ & s \geq 0. \end{aligned} \tag{6}$$

This problem (6) is very close to (3) with one difference that the variable z is free. We next express the variable z as the difference of two nonnegative variables $z = z^+ - z^-$ with $z^+ \geq 0, z^- \geq 0$. Then, the problem (6) is equivalent to

$$\begin{aligned} \min_{z^+, z^-, s} \quad & e^{\mathsf{T}}(z^+ - z^-) \\ \text{subject to} \quad & F(z^+ - z^-) - s = -g, \\ & z^+ \geq 0, z^- \geq 0, s \geq 0. \end{aligned} \tag{7}$$

The problem (7) is equivalent to the standard-form LP (3) by choosing

$$c = \begin{bmatrix} e \\ -e \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} F & -F & -I \end{bmatrix}, \quad b = -g, \quad x = \begin{bmatrix} z^+ \\ z^- \\ s \end{bmatrix}.$$

Very often, we also see a general linear program of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Ax = b \\ & Fx \leq h, \end{aligned} \tag{8}$$

where $A \in \mathbb{R}^{m \times n}, F \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^m, h \in \mathbb{R}^p, c \in \mathbb{R}^n$ are problem data. Similar to the conversion process above, we can transform a general LP (8) into the standard form LP (3) or the inequality-form LP (4). In all different forms of LPs (3), (4) or (8), it is clear that their feasible sets are *polyhedra* (see **??**), and thus we are optimizing a linear function over a polyhedron.

2.2 An LP example

LPs have a huge number of applications in many areas of applied sciences, engineering, and economics; see examples from e.g., [1, Chapter 1.3] and [3, Chapter 4.3]. We here discuss a simple example of LPs that has attracted extensive attention in machine learning and signal processing communities.

Consider the following optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \|x\|_1 \\ \text{subject to} \quad & Mx = d, \end{aligned} \tag{9}$$

where $M \in \mathbb{R}^{m \times n}$ and $d \in \mathbb{R}^m$ are problem data, and $m < n$. In (9), the linear equation $Mx = d$ is under-determined, and we aim to find a solution to $Mx = d$ that has the smallest l_1 norm. Recall that $\|x\|_1 = \sum_{i=1}^n |x_i|$.

Problem (9) is not an LP in its current form, since the cost function is not linear (instead, it is a piece-wise linear function). We shall see that (9) can be equivalently transformed into an LP by introducing new variables. The following result is a simple fact about the l_1 norm.

Lemma 5.1. *Let $x \in \mathbb{R}^n$. Its l_1 norm can be computed as*

$$\begin{aligned} \|x\|_1 = \min_{y \in \mathbb{R}^n} \quad & \sum_{i=1}^n y_i \\ \text{subjectto} \quad & y + x \geq 0, y - x \geq 0. \end{aligned}$$

Proof. First, we have $y \geq -x, y \geq x$, thus $|x_i| = \max(x_i, -x_i) \leq y_i$. We then have $\|x\|_1 \leq \sum_{i=1}^n y_i$ for any feasible y. On the other hand, the equality can be achieved by choosing $y_i = |x_i|, i = 1, \ldots, n$. □

By Lemma 5.1, we claim that (4) is equivalent to

$$\begin{aligned}\min_{x,y\in\mathbb{R}^n} \quad & \sum_{i=1}^n y_i \\ \text{subject to} \quad & Mx = d, \\ & y + x \geq 0, y - x \geq 0.\end{aligned} \tag{10}$$

Any solution of (10) can be converted to a solution of (9), and vice versa. Problem (10) is similar to an LP in the form (3), but it is not exactly in the form of (3). Let us define two new variables, $u = y + x$ and $v = y - x$, then (10) can be written as

$$\begin{aligned}\min_{x,y,u,v\in\mathbb{R}^n} \quad & \sum_{i=1}^n y_i \\ \text{subject to} \quad & Mx = d, \\ & u = y + x, v = y - x \\ & u \geq 0, v \geq 0.\end{aligned}$$

This problem is almost in the form of (3) with a small difference that x, y are not constrained to be nonnegative. In this case, we can eliminate x, y since $x = \frac{1}{2}(u - v)$, $y = \frac{1}{2}(u + v)$. Then, the problem becomes

$$\begin{aligned}\min_{u,v\in\mathbb{R}^n} \quad & \frac{1}{2}\sum_{i=1}^n (u_i + v_i) \\ \text{subject to} \quad & M(u - v) = 2d, \\ & u \geq 0, v \geq 0,\end{aligned}$$

which is an LP in the standard form (3) with matrices

$$A = \begin{bmatrix} M & -M \end{bmatrix} \in \mathbb{R}^{m\times 2n}, \qquad b = 2d \in \mathbb{R}^m, \qquad c = \frac{1}{2}\begin{bmatrix} 1, 1, \ldots, 1 \end{bmatrix} \in \mathbb{R}^{2n}.$$

The conversion from a problem (10) to its standard-form LP can be tedious. With some abuse of terminology, it is common to refer to a problem in (10) as an LP. However, note that the conversion from (9) to (10) is less obvious and cannot be generally taken for granted.

3 Semidefinite programming

3.1 Standard and inequality forms

A semidefinite program (SDP) is a conic optimization problem over the positive semidefinite cone $\mathbb{S}^n_+$. The standard-form SDP is a problem of the form

$$\begin{aligned}\min_{X} \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i, i = 1, \ldots, m \\ & X \in \mathbb{S}^n_+,\end{aligned} \tag{11}$$

where $C \in \mathbb{S}^n$ and $A_i \in \mathbb{S}^n, i = 1, \ldots, m$, $b_i \in \mathbb{R}, i = 1, \ldots, m$ are problem data, and the optimization variable is a matrix is $X \in \mathbb{S}^n$. Very often, we also use a linear map $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$, defined as

$$\mathcal{A}(X) = \begin{bmatrix} \langle A_1, X\rangle \\ \vdots \\ \langle A_m, X\rangle \end{bmatrix}$$

Then the affine constraints in (11) can also be written as $\mathcal{A}(X) = b$.

Example 5.2. *The following optimization problem is an SDP*

$$\begin{aligned} \min_{X} \quad & 2x_{11} + 2x_{12} \\ \textit{subject to} \quad & x_{11} + x_{22} = 1 \\ & \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \succeq 0. \end{aligned} \tag{12}$$

This SDP is in the standard-form (11) *with problem data* $C = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, b_1 = 1$ *and the conic constraint* $X \in K = \mathbb{S}^2_+$. □

The *inequality-form SDP* is a problem with one linear matrix inequality (LMI)

$$\begin{aligned} \min_{x \in \mathbb{R}^k} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & A_0 + x_1 A_1 + \cdots + x_k A_k \succeq 0, \end{aligned} \tag{13}$$

where $A_0, A_1, \ldots, A_k \in \mathbb{S}^n$ and $c \in \mathbb{R}^k$ are problem data. Similar to the LPs, the standard-form SDP (11) and the inequality-form SDP (13) can be converted equivalently from each other: to go from (11) to (13), we eliminate the affine constraints $\mathcal{A}(X) = b$ by finding a special solution and the basis of the nullspace of the linear mapping; to go from (13) to (11), we can add a slack variable $A_0 + x_1 A_1 + \cdots + x_k A_k = Z \succeq 0$ and split $x_i = x_i^+ - x_i^-$ with $x_i^+ \geq 0, x_i^- \geq 0, i = 1, \ldots, k$ (note that $\mathbb{R}^1_+ = \mathbb{S}^1_+$).

Example 5.3. *Consider the following problem*

$$\begin{aligned} \min_{x_1, x_2} \quad & 2x_1 + 2x_2 \\ \textit{subject to} \quad & \begin{bmatrix} x_1 & x_2 \\ x_2 & 1 - x_1 \end{bmatrix} \succeq 0. \end{aligned} \tag{14}$$

It is clear that (14) *is equivalent to* (12)*. Meanwhile, this problem is an inequality-form SDP* (13)*, where*

$$c = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \; A_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \; A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \; A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

□

A general-form SDP is an optimization problem of the form

$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & F_0 + x_1 F_1 + \ldots + x_k F_k \preceq 0, \end{aligned} \tag{15}$$

where $F_i \in \mathbb{S}^n, i = 0, \ldots, k$, $A \in \mathbb{R}^{m \times k}, b \in \mathbb{R}^m$ and $c \in \mathbb{R}^k$ are problem data.

Recall that a polyhedron is a set defined by finitely many linear inequalities (note that an equality $c_j^\mathsf{T} x = d_j$ is the same as two inequalities $c_j^\mathsf{T} x - d_j \leq 0, d_j - c_j^\mathsf{T} x \leq 0$) and the feasible region of LPs are polyhedra. Similarly, we define a *spectrahedron* as a set defined by finitely many linear matrix inequalities (LMIs). Note that several LMIs

$$F^{(i)}(x) = F_0^{(i)} + x_1 F_1^{(i)} + \ldots + x_k F_k^{(i)} \preceq 0, i = 1, \ldots, q$$

are equivalent to a single and large block-diagonal LMI as follows

$$\operatorname{diag}(F^{(1)}(x), \ldots, F^{(q)}(x)) \preceq 0.$$

Definition 5.1 (Spectrahedra). *A set $C \subseteq \mathbb{R}^k$ is called a spectrahedron if it has the form*

$$C = \left\{ x \in \mathbb{R}^k \mid A_0 + x_1 A_1 + \cdots + x_m A_m \succeq 0 \right\}, \tag{16}$$

for some matrices $A_i \in \mathbb{S}^n, i = 0, \ldots, m$.

Geometrically, a spectrahedron is an intersection of the positive semidefinite cone and an affine subspace $\{A_0 + x_1 A_1 + \cdots + x_m A_m \mid x \in \mathbb{R}^m\}$, and is thus a closed convex set (on the other hand, a spectrahedron can also be viewed as the pre-image of the positive semidefinite cone under the affine function $A(x) = A_0 + x_1 A_1 + \cdots + x_m A_m$). When the matrices $A_i, i = 0, \ldots, m$ in (16) are all diagonal, then the set C becomes a polyhedron. However, spectrahedra are much more general than polyhedra.

Example 5.4. *The unit disk $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ is not polyhedral, but it is a spectradedron. Indeed, we have*

$$D = \left\{ (x, y) \in \mathbb{R}^2 \mid \begin{bmatrix} 1 - x & y \\ y & 1 + x \end{bmatrix} \succeq 0 \right\}.$$

Also, the feasible region of the SDP (14) *is a closed disk*

$$\left\{ (x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - \frac{1}{2})^2 + x_2 \leq \frac{1}{4} \right\},$$

which not a polyhedron. □

In all different forms of SDPs (11), (13) or (15), it is clear that their feasible sets are *spectrahedra*. Geometrically, an SDP is a linear optimization problem over a spectrahedron. Any spectradedron is a closed convex set, but deciding which convex sets are spectradra (i.e., representing the convex set using an LMI of the form (16)) is an open research question. There is no known simple necessary and sufficient conditions (some nontrivial necessary conditions are known). We refer the interested reader to [2, Chapter 6] for details.

3.2 An SDP example

Semidefinite optimization is one central problem in this course. We will review more properties of SDPs (and the general conic programs) in **??** and discuss a wide range of applications in **??**.

We here discuss a symmetric matrix completion problem. This problem can be considered as a matrix analogue to (9): We observe certain entries of an unknown symmetric matrix and our goal is to recover the symmetric matrix with the smallest nuclear norm. Recall that the nuclear norm of a matrix $A \in \mathbb{R}^{m \times n}$ is the sum of its singular values $\|X\|_{\text{nuc}} = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)$. For symmetric matrices, the nuclear norm becomes the sum of the absolute values of its eigenvalues

$$\|X\|_{\text{nuc}} = \sum_{i=1}^{n} |\lambda_i(X)|,$$

where $\lambda_i(X)$ is the i-th eigenvalue of X. It can be interpreted as the l_1 norm of the eigenvalues of X (indeed, given a matrix $A \in \mathbb{R}^{m \times n}$, the nuclear norm is the l_1 norm of the singular eigenvalues, the spectral norm is the l_∞ norm of the singular eigenvalues, and the Frobenius norm is the l_2 norm of the singular eigenvalues).

Let $\mathcal{E} \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\}$ be the set of entries that we observe and $M_{ij}, (i, j) \in \mathcal{E}$ be the observed values. Then, we aim to solve the following problem

$$\begin{aligned} \min_{X \in \mathbb{S}^n} \quad & \|X\|_{\text{nuc}} \\ \text{subject to} \quad & X_{ij} = M_{ij}, \ \forall (i, j) \in \mathcal{E}. \end{aligned} \tag{17}$$

For the nuclear norm of symmetric matrices, we have the following result.

Lemma 5.2. *Let $X \in \mathbb{S}^n$. Its nuclear norm can be computed as follows*

$$\begin{aligned} \|X\|_{\mathrm{nuc}} = \min_{Y \in \mathbb{S}^n} \quad & \mathit{trace}(Y) \\ \text{subjectto} \quad & Y + X \succeq 0, Y - X \succeq 0. \end{aligned} \tag{18}$$

Proof. Similar to Lemma 5.1, it suffices to prove the following two statements:

- Any Y in the feasible region of (18) satisfies $\|X\|_{\mathrm{nuc}} \le \mathrm{trace}(Y)$.
- The equality can be achieved by a feasible Y.

Let us first show statement 2: denote the eigendecomposition of X and choose Y as

$$X = \sum_{i=1}^n \lambda_i v_i v_i^\mathsf{T}, \qquad Y = \sum_{i=1}^n |\lambda_i| v_i v_i^\mathsf{T}.$$

It is easy to see $\|X\|_{\mathrm{nuc}} = \mathrm{trace}(Y)$. In addition, we have

$$Y \pm X = \sum_{i=1}^n (|\lambda_i| \pm \lambda_i) v_i v_i^\mathsf{T} \succeq 0,$$

since $|\lambda_i| \pm \lambda_i \ge 0$. Thus, Y is in the feasible region of (18).

For statement 1, we denote

$$P^+ = \sum_{\lambda_i \ge 0} v_i v_i^\mathsf{T}, \qquad P^- = \sum_{\lambda_i < 0} v_i v_i^\mathsf{T},$$

which satisfies $P^+ + P^- = I$ (this is because $V = [v_1, \dots, v_n]$ is orthonormal, $VV^\mathsf{T} = I$) and $\mathrm{trace}(X(P^+ - P^-)) = \|X\|_{\mathrm{nuc}}$. Since $Y - X \succeq 0$ and $P^+ \succeq 0$, we have $\mathrm{trace}((Y - X)P^+) \ge 0$. Similarly, we have $\mathrm{trace}((Y + X)P^-) \ge 0$. This leads to

$$\begin{aligned} 0 \; &\le \mathrm{trace}((Y - X)P^+) + \mathrm{trace}((Y + X)P^-) \\ &= \mathrm{trace}((Y(P^+ + P^-) - X(P^+ - P^-)) \\ &= \mathrm{trace}(Y) - \|X\|_{\mathrm{nuc}}. \end{aligned}$$

This completes the proof. □

We can now claim that (17) can be equivalently formulated into an SDP below

$$\begin{aligned} \min_{Y, X \in \mathbb{S}^n} \quad & \mathrm{trace}(Y) \\ \text{subject to} \quad & X_{ij} = M_{ij}, \ \forall (i,j) \in \mathcal{E}, \\ & Y + X \succeq 0, \ Y - X \succeq 0. \end{aligned} \tag{19}$$

We can further introduce $U = Y + X$, $V = Y - X$ and then eliminate X and Y, leading to the *standard-form* SDP

$$\begin{aligned} \min_{U, V \in \mathbb{S}^n} \quad & \frac{1}{2}(\mathrm{trace}(U) + \mathrm{trace} V) \\ \text{subject to} \quad & U_{ij} + V_{ij} = 2M_{ij}, \ \forall (i,j) \in \mathcal{E}, \\ & U \succeq 0, \ V \succeq 0. \end{aligned}$$

It is common to refer to a problem in (19) as an SDP. We note that the conversion from (17) to (19) is not obvious and often requires non-trivial analysis. As you will see in HWs, the nuclear norm of a general matrix $A \in \mathbb{R}^{m \times n}$ can be computed by solving an SDP; see [5] for more details.

3.3 Solving semidefinite programs

There are efficient polynomial-time algorithms to solve the class of semidefinite optimization problems (e.g., interior-point algorithms); see [2, Section 2.3] for a summary of software implementation. We will not discuss these algorithms in this course. These algorithms can very reliably solve (11) with n up to a few hundreds and m up to a few thousands on a personal computer. Larger instances may be solved by exploiting structures (see, e.g., [6, Section 3.5] for a summary of algorithm implementation).

We recommend the MATLAB packages YALMIP (`https://yalmip.github.io/`) or CVX (`http://cvxr.com/cvx/`) (implementations in other languages, such as python, Julia, and R, are also available.) which allows us to use the natural description of SDPs and call another SDP solver to get a solution.

Notes

The preparation of this lecture was based on [4, Lectures 3 & 4]. Further reading for this lecture can refer to [2, Chapter 2].

References

[1] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[2] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2012.

[3] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[4] Hamza Fawzi. Topics in Convex Optimisation, Michaelmas 2018.

[5] Benjamin Recht, Maryam Fazel, and Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[6] Yang Zheng, Giovanni Fantuzzi, and Antonis Papachristodoulou. Chordal and factor-width decompositions for scalable semidefinite and polynomial optimization. *Annual Reviews in Control*, 2021.