CO 471 - Semi-definite Optimization

Typeset by Kevin Matthews

Spring 2014 (This incomplete draft was produced June 16, 2014.)

1 Introduction

REMARK **1.** Unfortunately, we take $\mathbb{N} = \{1, 2, \ldots\}$ in these notes. We also use $[n] := \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$.

DEFINITION **2.** A *linear optimization problem*, or a *Linear Programming (LP) program*, is a problem of minimizing or maximizing a linear function of finitely many real valued variables subject to finitely many linear equations and/or inequalities on those variables.

DEFINITION **3.** A *Semi-definite Programming (SDP) problem* is a problem of minimizing or maximizing a linear function of finitely many symmetric matrix variables with real entries subject to finitely many linear equations and linear inequalities on these variables and subject to positive semi-definiteness constraints on some of them.

DEFINITION **4.** Let $n \in \mathbb{N}$. We let $\mathbb{R}^{n \times n}$ denote the set of all n-by-n matrices with entries from $\mathbb{R}$. That is,

$$\mathbb{R}^{n \times n} := \left\{ \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{bmatrix} ; x_{i,j} \in \mathbb{R} \text{ for all } (i,j) \in [n]^2 \right\}.$$

DEFINITION **5.** Let $n \in \mathbb{N}$ and let $X \in \mathbb{R}^{n \times n}$. The *transpose* of X is given by

$${X^{\mathsf{T}}}_{i,j} := X_{j,i}.$$

DEFINITION **6.** Let $n \in \mathbb{N}$ and let $X \in \mathbb{R}^{n \times n}$. We say X is *symmetric* if $X^{\mathsf{T}} = X$.

DEFINITION **7.** Let $n \in \mathbb{N}$ and let $X \in \mathbb{R}^{n \times n}$. The *trace* of X is given by

$$\mathrm{Tr}(X) := \sum_{i=1}^{n} X_{i,i}.$$

Definition 8. Let $n \in \mathbb{N}$ and let $X, Y \in \mathbb{R}^{n \times n}$. The *inner product* of X and Y is given by

$$\langle X, Y \rangle := \mathrm{Tr}(X^T Y).$$

Proposition 9. *Let $n \in \mathbb{N}$ and let $X, Y \in \mathbb{R}^{n \times n}$. Then*

$$\langle X, Y \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,j} Y_{i,j}.$$

Proof. Simply note

$$\langle X, Y \rangle = \mathrm{Tr}(X^T Y) = \sum_{i=1}^{n} (X^T Y)_{i,i} = \sum_{i=1}^{n} \sum_{j=1}^{n} {X^T}_{i,j} Y_{j,i} = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{j,i} Y_{j,i}.$$

∎

Proposition 10. *Let $n, m \in \mathbb{N}$, let $X \in \mathbb{R}^{n \times m}$, and let $Y \in \mathbb{R}^{m \times n}$ Then*

$$\mathrm{Tr}(XY) = \mathrm{Tr}(YX).$$

Proof. We calculate

$$\mathrm{Tr}(XY) = \sum_{i=1}^{n} (XY)_{i,i} = \sum_{i=1}^{n} \sum_{j=1}^{m} X_{i,j} Y_{j,i} = \sum_{j=1}^{m} \sum_{i=1}^{n} Y_{j,i} X_{i,j} = \sum_{j=1}^{m} (YX)_{j,j} = \mathrm{Tr}(YX).$$

∎

Proposition 11. *Let $n \in \mathbb{N}$ and let $X, P \in \mathbb{R}^{n \times n}$. If P is non-singular (that is, $\det(P) \neq 0$) then $\mathrm{Tr}(PXP^{-1}) = \mathrm{Tr}(X)$.*

Proof. By the previous proposition, we calculate

$$\mathrm{Tr}(PXP^{-1}) = \mathrm{Tr}(XP^{-1}P) = \mathrm{Tr}(XI) = \mathrm{Tr}(X).$$

∎

Definition 12. Let $n \in \mathbb{N}$ and let $X \in \mathbb{R}^{n \times n}$. Note that $\det(X - \lambda I)$ is a polynomial in λ of degree n. By the fundamental theorem of algebra, it has n roots (possibly repeated) over $\mathbb{C}$. These roots are called the *eigenvalues* of X.

Proposition 13. *Let $n \in \mathbb{N}$ and let $A, B \in \mathbb{R}^{n \times n}$. If A and B are similar then they have the same eigenvalues.*

Proof. As A and B are similar, there is an invertible matrix $P \in \mathbb{R}^{n \times n}$ such that $A = PBP^{-1}$. Then

$$\det(A - \lambda I) = \det\left(PBP^{-1} - \lambda PIP^{-1}\right) = \det\left(P(B - \lambda I)P^{-1}\right) = \det(P)\det(B - \lambda I)\det(P)^{-1} = \det(B - \lambda I).$$

∎

DEFINITION 14. Let $n \in \mathbb{N}$. We let $\mathbb{S}^n$ denote the set of n-by-n symmetric matrices with entries in $\mathbb{R}$. That is,

$$\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n}; X = X^\mathsf{T}\}.$$

PROPOSITION 15. *Let $X \in \mathbb{S}^n$. Then the eigenvalues of X are all real numbers.*

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of X. This means that there exists a non-zero eigenvector v such that $Xv = \lambda v$. Note that since X is a real symmetric matrix, $X = \overline{X}$ and $X^\mathsf{T} = X$. It follows that $X\overline{v} = \overline{\lambda}\overline{v}$. So we get

$$\lambda \overline{v}^\mathsf{T} v = \overline{v}^\mathsf{T} \lambda v = \overline{v}^\mathsf{T} X v = (v^\mathsf{T} X \overline{v})^\mathsf{T} = (v^\mathsf{T} \overline{\lambda}\overline{v})^\mathsf{T} = \overline{\lambda}\overline{v}^\mathsf{T} v.$$

But

$$\overline{v}^\mathsf{T} v = \sum_{i=1}^{n} \overline{v}_i v_i = \langle v, v\rangle > 0$$

because v is non-zero.

The first equation gives us $(\lambda - \overline{\lambda})\overline{v}^\mathsf{T} v = 0$. Combining this with the second yields $\lambda = \overline{\lambda}$, or equivalently, $\lambda \in \mathbb{R}$. ∎

DEFINITION 16. Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. As the n eigenvalues of X are real numbers, we can order them. We let $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$ denote the eigenvalues of X in decreasing order. We also define

$$\lambda(X) := \begin{bmatrix} \lambda_1(X) \\ \vdots \\ \lambda_n(X) \end{bmatrix}.$$

DEFINITION 17. Let $n \in \mathbb{N}$ and let $X \in \mathbb{R}^{n\times n}$. We say X is *orthogonal* if $X^\mathsf{T} X = I$.

DEFINITION 18. Let $n \in \mathbb{N}$. We define $\mathrm{Diag} : \mathbb{R}^n \to \mathbb{S}^n$ by

$$\mathrm{Diag}(x) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}$$

for all $x \in \mathbb{R}^n$.

THEOREM 19. [THEOREM 1.8] *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. Then there exists an orthonormal matrix $Q \in \mathbb{R}^{n\times n}$ such that*

$$X = Q\,\mathrm{Diag}(\lambda(X))Q^\mathsf{T}. \tag{1}$$

Proof. Exercise. ∎

Definition 20. Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. For a matrix Q satisfying the previous theorem, the expression (1) is known as a *Schur decomposition/eigenvalue decomposition/spectral decomposition* of X.

Definition 21. Let $n \in \mathbb{N}$, let $X \in \mathbb{S}^n$, and let Q be a Schur decomposition of X. Then the columns $q^{(1)}, \dots, q^{(n)}$ of Q are the *eigenvectors* of X corresponding to the eigenvalues $\lambda_1(X), \dots, \lambda_n(X)$ respectively.

Definition 22. Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. We say X is *positive semi-definite* if for all $h \in \mathbb{R}^n$ we have $h^\mathsf{T} X h \geq 0$. We let

$$\mathbb{S}^n_+ := \{X \in \mathbb{S}^n ; x \text{ is positive semi-definite}\}.$$

Proposition 23. *Let* $n \in \mathbb{N}$ *and let* $X \in \mathbb{S}^n$. *Then* X *is positive semi-definite if and only if* $\lambda(X) \geq 0$.

Proof. Suppose first that X is positive semi-definite. Since the eigenvalues $\lambda(X)$ are ordered, it suffices to prove that $\lambda_n(X) \geq 0$.

Let Q be an orthonormal matrix which gives rise to a Schur decomposition of X. Consider $Qe_n \in \mathbb{R}^n$. Since X is positive semi-definite, we know $(Qe_n)^\mathsf{T} X Q e_n \geq 0$. But we can calculate

$$\begin{aligned}
(Qe_n)^\mathsf{T} X Q e_n &= {e_n}^\mathsf{T} Q^\mathsf{T} Q \operatorname{Diag}(\lambda(X)) Q^\mathsf{T} Q e_n \\
&= {e_n}^\mathsf{T} I \operatorname{Diag}(\lambda(X)) I e_n \\
&= {e_n}^\mathsf{T} \operatorname{Diag}(\lambda(X)) e_n \\
&= \sum_{i=1}^n \sum_{j=1}^n \delta_{i,n} \operatorname{Diag}(\lambda(X))_{i,j} \delta_{j,n} \\
&= \delta_{n,n} \operatorname{Diag}(\lambda(X))_{n,n} \delta_{n,n} \\
&= \lambda_n.
\end{aligned}$$

Thus $\lambda_n \geq 0$ as required.

Conversely, suppose that $\lambda(X) \geq 0$. Let $h \in \mathbb{R}^n$. We must show $h^\mathsf{T} X h \geq 0$. Again, let Q be an orthonormal matrix which gives rise to a Schur decomposition of X. Well, note

$$\begin{aligned}
h^\mathsf{T} X h &= h^\mathsf{T} Q \operatorname{Diag}(\lambda(X)) Q^\mathsf{T} h \\
&= (Q^\mathsf{T} h)^\mathsf{T} \operatorname{Diag}(\lambda(X)) (Q^\mathsf{T} h) \\
&= \sum_{i=1}^n \sum_{j=1}^n (Q^\mathsf{T} h)_i \operatorname{Diag}(\lambda(X))_{i,j} (Q^\mathsf{T} h)_j \\
&= \sum_{i=1}^n \lambda_i(X) (Q^\mathsf{T} h)_i^2
\end{aligned}$$

Since $\lambda(X) \geq 0$ we know $\lambda_i(X) \geq 0$, and clearly $(Q^\mathsf{T} h)_i^2 \geq 0$. Thus $h^\mathsf{T} X h \geq 0$. ∎

Remark 24. For $X \in \mathbb{S}^n$ and $h \in \mathbb{R}^n$,

$$\langle hh^\top, X\rangle = \mathrm{Tr}((hh^\top)^\top X) = \mathrm{Tr}(hh^\top X) = \mathrm{Tr}(h^\top Xh) = h^\top Xh.$$

Hence a constraint of the form, for all $h \in \mathbb{R}^n$ $h^\top Xh \geq 0$, is actually infinitely many linear constraints on X (of a special type of course).

Definition 25. Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. We say X is *positive definite* if for all $h \in \mathbb{R}^n \setminus \{0\}$ we have $h^\top Xh > 0$. We let

$$\mathbb{S}^n_{++} := \{X \in \mathbb{S}^n; X \text{ is positive definite}\}.$$

Proposition 26. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. Then X is positive definite if and only if $\lambda(X) > 0$.*

Proof. Suppose first that X is positive definite. Since the eigenvalues $\lambda(X)$ are ordered, it suffices to prove that $\lambda_n(X) > 0$.

Let Q be an orthonormal matrix which gives rise to a Schur decomposition of X. Consider $Qe_n \in \mathbb{R}^n \setminus \{0\}$ (indeed, Q is invertible and $e_n \neq 0$). Since X is positive definite, we know $(Qe_n)^\top XQe_n > 0$. But one can calculate, exactly as in the analogous proposition above, that $(Qe_n)^\top XQe_n = \lambda_n$. Thus $\lambda_n > 0$ as required.

Conversely, suppose that $\lambda(X) > 0$. Let $h \in \mathbb{R}^n \setminus \{0\}$. We must show $h^\top Xh > 0$. Again, let Q be an orthonormal matrix which gives rise to a Schur decomposition of X. Similarly to the previous analogous proposition, we have

$$h^\top Xh = \sum_{i=1}^n \lambda_i(X)(Q^\top h)_i^2 = \langle \bar{h}, \bar{h}\rangle$$

where $\bar{h} \in \mathbb{R}^n$ is given by $\bar{h}_i := \sqrt{\lambda_i(X)}(Q^\top h)_i$ for each $i \in [n]$. Note that since $\sqrt{\lambda_i(X)} > 0$, $h \neq 0$, and Q is invertible, we have that $\bar{h}$ is non-zero. Thus $h^\top Xh = \langle \bar{h}, \bar{h}\rangle > 0$. ∎

Definition 27. Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n_+$. The *square-root* of X is

$$X^{\frac{1}{2}} := Q \begin{bmatrix} \sqrt{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n(X)} \end{bmatrix} Q^\top,$$

where Q gives rise to a Schur decomposition of X.

Proposition 28. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n_+$. Then the square-root of X is well-defined.*

Proof. We need only show that the definition is independent of Q to show that $X^{\frac{1}{2}}$ is well-defined. Suppose $Q, R \in \mathbb{R}^{n\times n}$ both give rise to a Schur decomposition of X. Exercise. ∎

Proposition 29. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. If $X \in \mathbb{S}^n_+$ then $X^{\frac{1}{2}}$ is symmetric and $\lambda\left(X^{\frac{1}{2}}\right) = \sqrt{\lambda(X)}$, and if $X \in \mathbb{S}^n_{++}$ then X^{-1} is symmetric and $\lambda(X^{-1}) = \lambda(X)^{-1}$.*

Proof. Let Q give rise to a Schur decomposition of X.

Suppose $X \in \mathbb{S}^n_+$. One can compute directly that $\left(X^{\frac{1}{2}}\right)^\top = X^{\frac{1}{2}}$, so $X^{\frac{1}{2}}$ is symmetric. Also note that Q is a witness to the similarity of $X^{\frac{1}{2}}$ and $\sqrt{\mathrm{Diag}(X)}$, as is easily seen via the definition of square-root (recall $Q^\top = Q^{-1}$). Hence by proposition (13), $\lambda\left(X^{\frac{1}{2}}\right) = \lambda\left(\sqrt{\mathrm{Diag}(X)}\right) = \sqrt{\lambda(X)}$.

Suppose next $X \in \mathbb{S}^n_{++}$. Note that X is invertible since $\det(X) = \prod_{i=1}^n \lambda_i(X) \neq 0$. Now observe

$$X^{-1} = \left(Q\,\mathrm{Diag}(\lambda(X))Q^\top\right)^{-1} = Q\left(\mathrm{Diag}(\lambda(X))\right)^{-1}Q^\top = Q\begin{bmatrix} \frac{1}{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n(X)} \end{bmatrix} Q^\top.$$

It follows immediately by calculating the transpose of X^{-1} from this that X^{-1} is symmetric. Moreover, proposition (13) shows that $\lambda(X^{-1}) = \lambda(X)^{-1}$. ∎

Proposition 30. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. If $X \in \mathbb{S}^n_+$ then $X^{\frac{1}{2}} \in \mathbb{S}^n_+$, and if $X \in \mathbb{S}^n_{++}$ then $X^{\frac{1}{2}} \in \mathbb{S}^n_{++}$. Moreover, if $X \in \mathbb{S}^n_{++}$ then $X^{-1} \in \mathbb{S}^n_{++}$.*

Proof. Let Q give rise to a Schur decomposition of X.

Suppose first that $X \in \mathbb{S}^n_+$. By the previous proposition, $\lambda\left(X^{\frac{1}{2}}\right) = \sqrt{\lambda(X)} \geq 0$. Thus $X^{\frac{1}{2}} \in \mathbb{S}^n_+$.

Now suppose that $X \in \mathbb{S}^n_{++}$. By the previous proposition, $\lambda\left(X^{\frac{1}{2}}\right) = \sqrt{\lambda(X)} > 0$. Thus $X^{\frac{1}{2}} \in \mathbb{S}^n_{++}$.

Finally, suppose $X \in \mathbb{S}^n_{++}$. By the previous proposition, $\lambda(X^{-1}) = \lambda(X)^{-1} > 0$. Thus $X^{-1} \in \mathbb{S}^n_{++}$. ∎

Remark 31. For $X \in \mathbb{S}^n_{++}$, one can check (using the next proposition) that $(X^{-1})^{\frac{1}{2}} = \left(X^{\frac{1}{2}}\right)^{-1}$. Hence we may write $X^{-\frac{1}{2}}$. It follows from the previous proposition that $X^{-\frac{1}{2}} \in \mathbb{S}^n_{++}$.

Proposition 32. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n_+$. Then $X^{\frac{1}{2}}X^{\frac{1}{2}} = X$.*

Proof. Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix giving rise to a Schur decomposition of X.

Then we calculate

$$\begin{aligned}
X^{\frac{1}{2}}X^{\frac{1}{2}} &= Q\begin{bmatrix} \sqrt{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n(X)} \end{bmatrix} Q^T Q \begin{bmatrix} \sqrt{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n(X)} \end{bmatrix} Q^T \\
&= Q\begin{bmatrix} \sqrt{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n(X)} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1(X)} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2(X)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n(X)} \end{bmatrix} Q^T \\
&= Q\begin{bmatrix} \lambda_1(X) & 0 & \cdots & 0 \\ 0 & \lambda_2(X) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n(X) \end{bmatrix} Q^T \\
&= X,
\end{aligned}$$

since Q is orthogonal (that is, $Q^TQ = I$) and by the defining equation of Q (that is, $X = Q\,\mathrm{Diag}(\lambda(X))Q^T$). ∎

Lemma 33. *Let* $n \in \mathbb{N}$, *let* $X \in \mathbb{S}^n_+$, *and let* $i \in [n]$. *If* $X_{i,i} = 0$ *then for any* $j \in [n]$ *we have* $X_{i,j} = X_{j,i} = 0$.

Proof. Let $j \in [n] \setminus \{i\}$ (the result is given for $j = i$). Let $\beta \in \mathbb{R}$. Note

$$0 \le (\beta e_i + e_j)^T X (\beta e_i + e_j) = 2X_{i,j}\beta + X_{j,j}$$

because X is positive semi-definite and because $X_{i,i} = 0$. If $X_{i,j}$ were non-zero, then we would be able to pick a particular β such that $2X_{i,j}\beta + X_{i,j} < 0$. But this isn't the case, so we must have $X_{i,j} = 0$. Finally, since X is symmetric, $X_{j,i} = X_{i,j} = 0$. ∎

Theorem 34. [Theorem 1.9: Cholesky Characterization of $\mathbb{S}^n_+$] *Let* $n \in \mathbb{N}$ *and let* $X \in \mathbb{S}^n$. *Then* $X \in \mathbb{S}^n_+$ *if and only if there exists a lower triangular matrix* $B \in \mathbb{R}^{n \times n}$ *such that* $X = BB^T$.

Proof. Suppose first that there exists a lower triangular matrix $B \in \mathbb{R}^{n\times n}$ such that $X = BB^T$. If $h \in \mathbb{R}^n$ then

$$h^TXh = h^TBB^Th = (B^Th)^T I (B^Th) \ge 0$$

because I is positive semi-definite and $B^Th \in \mathbb{R}^n$. Note that we did not use the fact that B is lower triangular.

Conversely, suppose that $X \in \mathbb{S}^n_+$. We will prove the claim by induction on n. Suppose $n = 1$. This means $X = \begin{bmatrix} X_{1,1} \end{bmatrix}$ and $\lambda(X) = \begin{bmatrix} X_{1,1} \end{bmatrix}$. Since X is positive semi-definite, $\lambda(X) \ge 0$, and so $X_{1,1} \geq 0$. Then we can take $B = \left[\sqrt{X_{1,1}}\right]$, which is trivially lower triangular, to get

$$BB^T = \left[\sqrt{X_{1,1}}\right]\left[\sqrt{X_{1,1}}\right]^T = \left[\sqrt{X_{1,1}}\right]\left[\sqrt{X_{1,1}}\right] = \left[X_{1,1}\right] = X.$$

Now suppose $n > 1$ and that the result holds for $n-1$. Pick $x \in \mathbb{R}^{n-1}$ and $\bar{X} \in \mathbb{R}^{n-1}$ so that

$$X = \begin{bmatrix} X_{1,1} & x^T \\ x & \bar{X} \end{bmatrix}.$$

Note that $X_{1,1} \geq 0$. Indeed, if $X_{1,1} < 0$ then

$$e_1^T X e_1 = X_{1,1} < 0,$$

contradicting that $X \in \mathbb{S}^n_+$. (Note that if $X_{1,1} < 0$, then we have a certificate $h \in \mathbb{R}^n$ such that $h^T X h < 0$ — in this case $h = e_1$.)

So we have $X_{1,1} \geq 0$. Also note that $\bar{X} \in \mathbb{S}^{n-1}_+$. Indeed, $\bar{X}$ is clearly symmetric, and if $h \in \mathbb{R}^{n-1}$ then

$$h^T \bar{X} h = \begin{bmatrix} 0 & h^T \end{bmatrix} \begin{bmatrix} X_{1,1} & x^T \\ x & \bar{X} \end{bmatrix} \begin{bmatrix} 0 \\ h \end{bmatrix} \geq 0$$

because $X \in \mathbb{S}^n_+$.

If $X_{1,1} = 0$, then by lemma (33) we know $x = 0$. By inductive assumption, there exists a lower triangular matrix $\bar{B} \in \mathbb{R}^{n-1 \times n-1}$ such that $\bar{X} = \bar{B}\bar{B}^T$. Taking

$$B := \begin{bmatrix} 0 & 0^T \\ 0 & \bar{B} \end{bmatrix}$$

we see

$$BB^T = \begin{bmatrix} 0 & 0^T \\ 0 & \bar{B}\bar{B}^T \end{bmatrix} = \begin{bmatrix} 0 & 0^T \\ 0 & \bar{X} \end{bmatrix} = X$$

as required.

So we may assume $X_{1,1} > 0$. Let $b := \frac{1}{\sqrt{X_{1,1}}} x$ and consider

$$\tilde{X} := \bar{X} - bb^T = \bar{X} - \tfrac{1}{X_{1,1}} xx^T.$$

Assume that $\tilde{X} \notin \mathbb{S}^{n-1}_+$. It is easy to see that $\tilde{X}$ is symmetric, so this means that there is an $\tilde{h} \in \mathbb{R}^{n-1}$ such that $\tilde{h}^T \tilde{X} \tilde{h} < 0$. Set $h_0 := -\frac{1}{X_{1,1}} x^T \tilde{h}$ and consider

$$h := \begin{bmatrix} h_0 \\ \tilde{h} \end{bmatrix} \in \mathbb{R}^n.$$

We compute

$$\begin{aligned} h^T X h &= -\tfrac{1}{X_{1,1}} \left(x^T \tilde{h}\right)^2 + \tilde{h}^T \bar{X} \tilde{h} \\ &= -\tfrac{1}{X_{1,1}} \left(x^T \tilde{h}\right)^2 + \tilde{h}^T \left(bb^T + \tilde{X}\right) \tilde{h} \\ &= -\tfrac{1}{X_{1,1}} \left(x^T \tilde{h}\right)^2 + \tilde{h}^T bb^T \tilde{h} + \tilde{h}^T \tilde{X} \tilde{h} \\ &= \tilde{h}^T \tilde{X} \tilde{h} < 0, \end{aligned}$$

giving a certificate of non-positive semi-definiteness.

So we may assume $\tilde{X} \in \mathbb{S}_+^{n-1}$. Thus by inductive assumption, there exists a lower triangular matrix $\tilde{B} \in \mathbb{R}^{n-1\times n-1}$ such that $\tilde{B}\tilde{B}^T = \tilde{X}$. Then we will take

$$B := \begin{bmatrix} \sqrt{X_{1,1}} & 0^T \\ b & \tilde{B} \end{bmatrix},$$

which is clearly lower triangular. Computing

$$BB^T = \begin{bmatrix} \sqrt{X_{1,1}} & 0^T \\ b & \tilde{B} \end{bmatrix} \begin{bmatrix} \sqrt{X_{1,1}} & b^T \\ 0 & \tilde{B}^T \end{bmatrix} = \begin{bmatrix} X_{1,1} & \sqrt{X_{1,1}}b^T \\ \sqrt{X_{1,1}}b & bb^T + \tilde{B}\tilde{B}^T \end{bmatrix} = \begin{bmatrix} X_{1,1} & x^T \\ x & \bar{X} \end{bmatrix} = X$$

yields the desired result. ∎

Definition 35. Let $n \in \mathbb{N}$, let $p \in [1, \infty)$, and let $h \in \mathbb{R}^n$. The p-*norm* of h is

$$\|h\|_p := \left(\sum_{j=1}^{n} |h_j|^p \right)^{\frac{1}{p}}.$$

The ∞-*norm* of h is

$$\|h\|_\infty := \max\{|h_j|;\, j \in [n]\}.$$

Definition 36. Let $n \in \mathbb{N}$, let $p \in [1, \infty]$, and let $X \in \mathbb{S}^n$. The *operator* p-*norm* of X is

$$\|X\|_p := \max\{\|Xh\|_p;\, h \in \mathbb{R}^n \text{ and } \|h\|_p = 1\}.$$

The *Frobenius-norm* of X is

$$\|X\|_F := \langle X, X\rangle^{\frac{1}{2}}.$$

Remark 37. Let $n \in \mathbb{N}$, let $p \in [1, \infty]$, and let $X \in \mathbb{S}^n$. Then

$$\|X\|_F = \langle X, X\rangle^{\frac{1}{2}} = \left(\mathrm{Tr}(X^TX)\right)^{\frac{1}{2}} = (\mathrm{Tr}(XX))^{\frac{1}{2}} = \|\lambda(X)\|_2.$$

To see this last equality, let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix which gives rise to a Schur decomposition of X, and observe

$$\begin{aligned}
\mathrm{Tr}(XX) &= \mathrm{Tr}\left(Q\,\mathrm{Diag}(\lambda(X))Q^{\mathsf{T}}Q\,\mathrm{Diag}(\lambda(X))Q^{\mathsf{T}}\right)\\
&= \mathrm{Tr}\left(Q\,\mathrm{Diag}(\lambda(X))^2Q^{\mathsf{T}}\right)\\
&= \mathrm{Tr}\left(\mathrm{Diag}(\lambda(X))^2Q^{\mathsf{T}}Q\right)\\
&= \mathrm{Tr}\left(\mathrm{Diag}(\lambda(X))^2\right)\\
&= \mathrm{Tr}\left(\begin{bmatrix}\lambda_1(X)^2 & 0 & \cdots & 0\\ 0 & \lambda_2(X)^2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n(X)^2\end{bmatrix}\right)\\
&= \sum_{i=1}^{n}\lambda_i(X)^2\\
&= \|\lambda(X)\|_2^2.
\end{aligned}$$

Definition 38. Let $n \in \mathbb{N}$, let $J \subseteq [n]$ be non-empty, and let $X \in \mathbb{R}^{n\times n}$. We define X_J to be the matrix X with each entries (i,j) where $(i,j) \in J^2$.

Example 39. For example,

$$\begin{bmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ 9 & 10 & \sqrt{2} & 12\\ 13 & 14 & \pi & 16\end{bmatrix}_{\{1,3\}} = \begin{bmatrix}1 & 3\\ 9 & \sqrt{2}\end{bmatrix}.$$

Lemma 40. *Let $n \in \mathbb{N}$, let $J \subseteq [n]$ be non-empty, and let $X \in \mathbb{S}^n_+$. Then $X_J \in \mathbb{S}^{\#J}_+$.*

Proof. Let $h \in \mathbb{R}^{\#J}$ (and assume it is indexed by J instead of $[\#J]$). Define $\bar{h} \in \mathbb{R}^n$ by filling in the indices not in J with zeros. Since $X \in \mathbb{S}^n_+$ and because $\bar{h}_t = 0$ for $t \notin J$, we get

$$h^{\mathsf{T}}X_Jh = \sum_{k\in J}\sum_{\ell\in J} h_k\,(X_J)_{k,\ell}\,h_\ell = \sum_{k\in J}\sum_{\ell\in J}\bar{h}_kX_{k,\ell}\bar{h}_\ell = \sum_{k=1}^{n}\sum_{\ell=1}^{n}\bar{h}_kX_{k,\ell}\bar{h}_\ell = \bar{h}^{\mathsf{T}}X\bar{h} \geq 0.$$

Hence $X_J \in \mathbb{S}^{\#J}$. ∎

Proposition 41. [Proposition 1.10] *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. Then the following are equivalent:*

(1) $X \in \mathbb{S}^n_+$;

(2) $\lambda(X) \geq 0$;

(3) There exists a $\mu \in \mathbb{R}^n_+$ *and* $h^{(1)},\ldots,h^{(n)} \in \mathbb{R}^n$ *such that* $X = \sum_{k=1}^{n}\mu_k h^{(k)}{h^{(k)}}^{\mathsf{T}}$;

(4) *There exists a* $B \in \mathbb{R}^{n\times n}$ *such that* $X = BB^\mathsf{T}$ *(and we may choose* B *to be lower triangular);*

(5) *For every non-empty* $J \subseteq [n]$, $\det(X_J) \geq 0$*;*

(6) *For all* $S \in \mathbb{S}^n_+$, $\langle S, X\rangle \geq 0$.

Proof. We already saw in proposition (23) that (1) and (2) are equivalent.

Next we show that (3) is equivalent to the first two. Suppose first that $X \in \mathbb{S}^n_+$. Consider $\lambda(X) \in \mathbb{R}^n_+$ (by (2)), and let Q give rise to a Schur decomposition of X. Let $q^{(1)}, \dots, q^{(n)}$ denote the columns of Q. Suppose $(i,j) \in [n]^2$. Then note

$$X_{i,j} = \left(Q\,\mathrm{Diag}(\lambda(X))Q^\mathsf{T}\right)_{i,j} = \sum_{k=1}^n \lambda_k(X) Q_{i,k} Q_{k,j},$$

while

$$\left(\sum_{k=1}^n \lambda(X)_k q^{(k)} {q^{(k)}}^\mathsf{T}\right)_{i,j} = \sum_{k=1}^n \lambda_k(X) Q_{i,k} Q_{k,j}.$$

Thus

$$X = \sum_{k=1}^m \lambda(X)_k q^{(k)} {q^{(k)}}^\mathsf{T}.$$

Conversely, suppose (3) is satisfied. Suppose $h \in \mathbb{R}^n$. Observe

$$h^\mathsf{T} X h = h^\mathsf{T}\left(\sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^\mathsf{T}\right) h = \sum_{k=1}^n \mu_k \left(h^\mathsf{T} h^{(k)}\right)\left(h^\mathsf{T} h^{(k)}\right)^\mathsf{T} \geq 0.$$

Indeed, we have seen many times that for $x^\mathsf{T} \in \mathbb{R}^n$ we have $xx^\mathsf{T} \geq 0$ (it is a sum of squares!) and we are given that $\mu_k \geq 0$.

We already saw in theorem (34) that (1) is equivalent to (4).

Now we show that (5) is equivalent to (1) through (4). Suppose first that (1) through (4) hold, and let $J \subseteq [n]$ be non-empty. By lemma (40), $X_J \in \mathbb{S}^{\#J}$. By (2), we get $\lambda(X_J) \geq 0$. Thus $\det(X_J) = \prod \lambda(X_J) \geq 0$.

Conversely, suppose that (5) holds. Exercise.

Finally, we show that (6) is equivalent to (1) up to (5). First, suppose that for all $S \in \mathbb{S}^n_+$ we have $\langle S, X\rangle \geq 0$. Suppose $h \in \mathbb{R}^n$. Then

$$h^\mathsf{T} X h = \mathrm{Tr}(h^\mathsf{T} X h) = \mathrm{Tr}(hh^\mathsf{T} X) = \left\langle hh^\mathsf{T}, X\right\rangle \geq 0$$

because $hh^\mathsf{T} \in \mathbb{S}^n_+$ (we have seen this many times). Thus $X \in \mathbb{S}^n_+$.

Conversely, suppose that $X \in \mathbb{S}^n_+$. Let $S \in \mathbb{S}^n_+$. By (3), there exists $\mu \in \mathbb{R}^n_+$ and $h^{(1)}, \ldots, h^{(n)} \in \mathbb{R}^n$ such that $S = \sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^\mathsf{T}$. So

$$\begin{aligned}
\langle S, X \rangle &= \operatorname{Tr}(S^\mathsf{T} X) \\
&= \operatorname{Tr}(SX) \\
&= \operatorname{Tr}\left(\sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^\mathsf{T} X\right) \\
&= \sum_{k=1}^n \mu_k \operatorname{Tr}\left(h^{(k)} {h^{(k)}}^\mathsf{T} X\right) \\
&= \sum_{k=1}^n \mu_k \operatorname{Tr}\left({h^{(k)}}^\mathsf{T} X h^{(k)}\right) \\
&= \sum_{k=1}^n \mu_k {h^{(k)}}^\mathsf{T} X h^{(k)} \geq 0
\end{aligned}$$

because $\mu \geq 0$ and $X \in \mathbb{S}^n_+$. ∎

Proposition 42. *Let $n \in \mathbb{N}$. Then $\operatorname{int}(\mathbb{S}^n_+) = \mathbb{S}^n_{++}$ (as subspaces of $\mathbb{S}^n$).*

Proof. Let $X \in \mathbb{S}^n_+ +$ and let $\epsilon > 0$. Suppose $S \in \mathbb{S}^n$. Exercise — how best to do this? ∎

Proposition 43. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. Then the following are equivalent:*

(1) $X \in \mathbb{S}^n_{++}$;

(2) $\lambda(X) > 0$;

(3) There exists a $\mu \in \mathbb{R}^n_{++}$ and linearly independent $\{h^{(1)}, \ldots, h^{(n)}\} \subseteq \mathbb{R}^n$ such that $X = \sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^\mathsf{T}$;

(4) There exists an invertible $B \in \mathbb{R}^{n\times n}$ such that $X = BB^\mathsf{T}$ (and we may choose B to be lower triangular);

(5) For all $k \in [n]$, $\det(X_{[k]}) > 0$;

(6) For all $S \in \mathbb{S}^n_+ \setminus \{0\}$, $\langle S, X \rangle > 0$;

(7) $X \in \mathbb{S}^n_+$ and $\operatorname{rank}(X) = n$.

Proof. We saw in proposition (26) that (1) and (2) are equivalent.

First we show that (1) and (2) are equivalent to (3). Suppose first that $X \in \mathbb{S}^n_{++}$. Let $Q \in \mathbb{R}^{n\times n}$ give rise to a Schur decomposition of X and let $q^{(1)}, \ldots, q^{(n)} \in \mathbb{R}^n$ denote the columns of Q. Note that since Q is invertible (because $QQ^\mathsf{T} = I$) we must have that $\{q^{(1)}, \ldots, q^{(n)}\}$ is linearly independent. Finally, note

$$X = Q \operatorname{Diag}(\lambda(X)) Q^\mathsf{T} = \sum_{k=1}^n \lambda_k(X) q^{(k)} {q^{(k)}}^\mathsf{T}.$$

Conversely, suppose that (3) holds. Let $h \in \mathbb{R}^n \setminus \{0\}$. Then

$$(2) \qquad h^T X h = h^T \left(\sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^T \right) h = \sum_{k=1}^n \mu_k \left({h^{(k)}}^T h \right)^T \left({h^{(k)}}^T h \right) = \sum_{k=1}^n \mu_k \left({h^{(k)}}^T h \right)^2 \geq 0$$

since $\mu > 0$. To show strict inequality, it suffices to show that one of the terms is positive. Assume, for a contradiction, that $\left({h^{(k)}}^T h \right)^2 = 0$ for each $k \in [n]$. This means $\langle h^{(k)}, h \rangle = {h^{(k)}}^T h = 0$ for each $k \in [n]$. But $\{h^{(1)}, \dots, h^{(n)}\}$ is linearly independent and hence a basis, so $\langle x, h \rangle = 0$ for all $x \in \mathbb{R}^n$. This implies $h = 0$. Contradiction. Thus one of the terms in equation (2) is positive as required and $X \in \mathbb{S}^n_{++}$.

Next we show the equivalence of (7) to (2). Suppose first that $X \in \mathbb{S}^n_{++}$. Clearly $S \in \mathbb{S}^n_+$. Moreover,

$$\det(X) = \prod_{k=1}^n \lambda_k(X) > 0$$

and hence X is invertible, or equivalently, $\operatorname{rank}(X) = n$.

Conversely, suppose that $X \in \mathbb{S}^n_+$ and $\operatorname{rank}(X)$. The former assumption shows that $\prod_{k=1}^n \lambda_k(X) \geq 0$. The latter says $\prod_{k=1}^n \lambda_k(X) = \det(X) \neq 0$, and so we conclude $\prod_{k=1}^n \lambda_k(X) > 0$. Hence $\lambda_k(X) \neq 0$ for all $k \in [n]$, and so since we already knew $\lambda_k(X) \geq 0$ we conclude $\lambda_k(X) > 0$ as required.

Next we show that (4) is equivalent to (1) and (7). Suppose first that $X \in \mathbb{S}^n_{++}$. By theorem (34), we get that there exists a lower triangular matrix $B \in \mathbb{R}^{n \times n}$ such that $X = BB^T$. But $\det(B)^2 = \det(B)\det(B^T) = \det(BB^T) = \det(X) \neq 0$ by (7), so B is invertible.

Conversely, suppose that (4) holds. Then we already know from the characterization of positive semi-definite matrices that $X \in \mathbb{S}^n_+$. But also note that $\det(X) = \det(BB^T) = \det(B)\det(B^T) = \det(B)^2 \neq 0$ and so $\operatorname{rank}(X) = n$. Thus (7) is satisfied.

Next we show that (6) is equivalent to (3). Suppose first that $X \in \mathbb{S}^n_{++}$. Let $S \in \mathbb{S}^n_{++}$. By (3), there exists $\mu \in \mathbb{R}^n_{++}$ and linearly independent (in particular, non-zero) $h^{(1)}, \dots, h^{(n)} \in \mathbb{R}^n$ such that $S = \sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^T$. So

$$\begin{aligned}
\langle S, X \rangle &= \operatorname{Tr}(S^T X) \\
&= \operatorname{Tr}(SX) \\
&= \operatorname{Tr}\left(\sum_{k=1}^n \mu_k h^{(k)} {h^{(k)}}^T X \right) \\
&= \sum_{k=1}^n \mu_k \operatorname{Tr}\left(h^{(k)} {h^{(k)}}^T X \right) \\
&= \sum_{k=1}^n \mu_k \operatorname{Tr}\left({h^{(k)}}^T X h^{(k)} \right) \\
&= \sum_{k=1}^n \mu_k {h^{(k)}}^T X h^{(k)} > 0
\end{aligned}$$

because $\mu > 0$ and $X \in \mathbb{S}^n_{++}$ (recall $h^{(k)}$ is non-zero for each $k \in [n]$).

Conversely, suppose that (6) holds. Suppose $h \in \mathbb{R}^n \setminus \{0\}$. Then

$$h^T X h = \mathrm{Tr}(h^T X h) = \mathrm{Tr}(hh^T X) = \langle hh^T, X \rangle > 0$$

because $hh^T \in \mathbb{S}^n_+ \setminus \{0\}$ (as $h \neq 0$). Thus $X \in \mathbb{S}^n_{++}$.

Finally we show that (5) is equivalent to the others. Exercise. ∎

Definition 44. Let $n \in \mathbb{N}$, let $j \in [n]$, and let $M \in \mathbb{R}^{n \times n}$. We define

$$B_j(M) := \left\{ \lambda \in \mathbb{C};\, |\lambda - M_{j,j}| \leq \sum_{i \neq j} |M_{j,i}| \right\}.$$

Theorem 45. [Gerschgorin Disc Theorem] *Let $n \in \mathbb{N}$, let $\lambda \in \mathbb{C}$, and let $M \in \mathbb{R}^{n \times n}$. If λ is an eigenvalue of M then*

$$\lambda \in \bigcup_{j=1}^{n} B_j(M).$$

Proof. Exercise. ∎

Definition 46. Let $n \in \mathbb{N}$ and let $M \in \mathbb{R}^{n \times n}$. We say M is *diagonally dominant* if for all $i \in [n]$ we have

$$X_{i,i} \geq \sum_{j \neq i} |X_{i,j}|.$$

We say M is *strictly diagonally dominant* if for all $i \in [n]$ we have

$$X_{i,i} > \sum_{j \neq i} |X_{i,j}|.$$

Corollary 47. *Let $n \in \mathbb{N}$ and let $X \in \mathbb{S}^n$. If X is diagonally dominant then $X \in \mathbb{S}^n_+$, and if X is strictly diagonally dominant then $X \in \mathbb{S}^n_{++}$.*

Proof. Suppose first that X is diagonally dominant and let $i \in [n]$. It follows immediately that $B_j(X) \subseteq \{z \in \mathbb{C}; \Re(z) \geq 0\}$ for each $j \in [n]$. By theorem (45), $\lambda_i(X) \in B_j(X)$ for some $j \in [n]$. Since X is symmetric, we know that $\lambda_i(X) \in \mathbb{R}$. Combining the above two observations we see $\lambda_i(X) = \Re(\lambda_i(X)) \geq 0$. Thus $\lambda(X) \geq 0$. By our many equivalences of positive semi-definiteness, $X \in \mathbb{S}^n_+$.

Next suppose that X is strictly diagonally dominant and let $i \in [n]$. This time we have $B_j(X) \subseteq \{z \in \mathbb{C}; \Re(z) > 0\}$ for each $j \in [n]$. By theorem (45), $\lambda_i(X) \in B_j(X)$ for some $j \in [n]$. Since X is symmetric we have $\lambda_i(X) \in \mathbb{R}$. Combining the above two observations we see $\lambda_i(X) = \Re(\lambda_i(X)) > 0$. Thus $\lambda(X) > 0$. By our many equivalences of positive definite matrices, $X \in \mathbb{S}^n_{++}$. ∎

COROLLARY **48.** *Let* $n \in \mathbb{N}$ *and let* $M \in \mathbb{S}^n$. *Then there exists a* $\mu_0 \in \mathbb{R}$ *such that* $X + \mu I \in \mathbb{S}^n_{++}$ *for all* $\mu \geq \mu_0$.

Proof. Let $\mu \in \mathbb{R}$. Note that

$$(X + \mu I)_{i,i} = X_{i,i} + \mu > \sum_{j \neq i} |X_{i,j}| = \sum_{j \neq i} \left|(X + \mu I)_{i,j}\right|$$

for sufficiently large μ. That is to say, $X + \mu I$ is strictly diagonally dominant for sufficiently large μ and hence, by the previous corollary, positive definite. ∎

DEFINITION **49.** Let $n \in \mathbb{N}$ and let $X, S \in \mathbb{S}^n$. We write $X \succcurlyeq S$ if $X - S \in \mathbb{S}^n_+$. We write $X \succ S$ if $X - S \in \mathbb{S}^n_{++}$.

DEFINITION **50.** Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$. We say K is a *convex cone* if for all $x, v \in K$ and all $\alpha \in \mathbb{R}_{++}$ we have $x + v \in K$ and $\alpha x \in K$.

DEFINITION **51.** Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a convex cone. We say K is *pointed* if it contains no lines.

DEFINITION **52.** Let $n, m \in \mathbb{N}$ and let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$. The *adjoint* of $\mathcal{A}$, denoted $\mathcal{A}^*$, is given by

$$\langle \mathcal{A}^*(y), X \rangle_{\mathbb{S}^n} := y^\mathsf{T} \mathcal{A}(x)$$

for all $y \in \mathbb{R}^m$ and all $X \in \mathbb{S}^n$.

REMARK **53.** In the situation above, there exist matrices $A_1, \ldots, A_m \in \mathbb{S}^n$ such that $(\mathcal{A}(X))_k = \langle A_k, X \rangle$ for each $k \in [m]$. But now we can see that for any $y \in \mathbb{R}^m$ and any $X \in \mathbb{S}^n$ we have

$$\left\langle \sum_{k=1}^m y_k A_k, X \right\rangle = \sum_{k=1}^m y_k \langle A_k, X \rangle = \sum_{k=1}^m y_k (\mathcal{A}(X))_k = y^\mathsf{T} \mathcal{A}(X).$$

Thus

$$\mathcal{A}^*(y) = \sum_{k=1}^m y_k A_k.$$

DEFINITION **54.** Let $n, m \in \mathbb{N}$, $C \in \mathbb{S}^n$, $b \in \mathbb{R}^m$, and let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear transformation. The *primal SDP in standard equality form*, denoted (P), is:

$$\begin{aligned} \inf \quad & \langle C, X \rangle \\ \mathcal{A}(X) &= b \\ X &\succcurlyeq 0 \end{aligned}$$

The *dual* SDP of (P), denoted (D), is:

$$\begin{aligned} \sup \quad & b^T y \\ \mathcal{A}^*(y) & \preccurlyeq C \end{aligned}$$

REMARK **55.** Continuing from the previous remark, we may rewrite the primal SDP as:

$$\begin{aligned} \inf \quad & \langle C, X \rangle \\ \langle A_i, X \rangle = b_i & \quad \text{for each } i \in [n] \\ X \in \mathbb{S}^n_+ & \end{aligned}$$

And the dual SDP can be written as:

$$\begin{aligned} \sup \quad & b^T y \\ \mathcal{A}^*(y) + S = C & \\ S \succcurlyeq 0 & \end{aligned}$$

DEFINITION **56.** Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a convex cone. The *dual cone* of K is

$$K^* := \{s \in \mathbb{E}; \langle s, x \rangle \geq 0 \ \forall x \in K\}.$$

PROPOSITION **57.** *Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a convex cone. Then K^* is a convex cone.*

Proof. Suppose $v, w \in K^*$ and $\alpha \in \mathbb{R}_{++}$. Suppose $x \in K$. Then

$$\langle v + w, x \rangle = \langle v, x \rangle + \langle w, x \rangle \geq 0 + 0 = 0$$

and

$$\langle \alpha v, x \rangle = \alpha \langle v, x \rangle \geq \langle v, x \rangle \geq 0.$$

Thus $v + w, \alpha v \in K^*$ and so K^* is a convex cone. ∎

REMARK **58.** The primal-dual pair is useful in more general settings. In fact, we can replace the convex cone $\mathbb{S}^n_+$ by any convex cone K. That is, given $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$, and a convex cone $K \subseteq \mathbb{R}^n$, we can look at

$$\begin{aligned} \inf \quad & \langle c, x \rangle \\ \mathcal{A}(x) & = b \\ x & \in K \end{aligned} \qquad\qquad \begin{aligned} \sup \quad & \langle b, y \rangle \\ \mathcal{A}^*(y) + s & = c \\ s & \in K^* \end{aligned}$$

Definition 59. Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a convex cone. The *automorphism group* of K is

$$\mathrm{Aut}(K) := \{T : \mathbb{E} \to \mathbb{E};\ T \text{ is an invertible linear transformation such that } T(K) = K\}.$$

Definition 60. Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a pointed, closed convex cone with non-empty interior. We say K is *self-dual* if there exists an inner product on $\mathbb{R}^n$ under which $K = K^*$. We say K is *homogeneous* if for every pair $x, \nu \in \mathrm{int}(K)$ there exists an automorphism $T \in \mathrm{Aut}(K)$ such that $T(x) = \nu$. We say K is *symmetric* if it is both self-dual and homogeneous.

Definition 61. Let $\mathbb{E}$ be a Euclidean space and let $K \subseteq \mathbb{E}$ be a pointed, closed convex cone with non-empty interior. A *ray* of K is a set of the form

$$\{\alpha x;\ \alpha \in \mathbb{R}_+\}$$

for some $x \in K \setminus \{0\}$.

Definition 62. Let $\mathbb{E}$ be a Euclidean space, let $K \subseteq \mathbb{E}$ be a pointed, closed convex cone with non-empty interior, and let R be a ray of K. We say R is *extreme* if for every pair of rays R_1, R_2 of K such that $R \subseteq R_1 + R_2$, we have $R \in \{R_1, R_2\}$.

Theorem 63. *Let $n \in \mathbb{N}$. Then $\mathbb{S}^n_+$ is a pointed, closed convex cone with non-empty interior which is symmetric (under the trace inner product). Moreover,*

$$\mathrm{ext}(\mathbb{S}^n_+) = \{hh^T;\ h \in \mathbb{R}^n,\ \|h\|_2 = 1\}.$$

Proof. First note that by our characterization of positive semi-definiteness, we saw that for a symmetric matrix $X \in \mathbb{S}^n$ we have $X \in \mathbb{S}^n_+$ if and only if $\langle X, S\rangle \geq 0$ for all $S \in \mathbb{S}^n_+$. That is to say,

$$\mathbb{S}^n_+ = \{X \in \mathbb{S}^n;\ \langle X, S\rangle \geq 0\ \forall S \in \mathbb{S}^n\} = \bigcap_{S \in \mathbb{S}^n} \{X \in \mathbb{S}^n;\ \langle X, S\rangle \geq 0\} \tag{3}$$

is the intersection of closed convex cones (check), and is therefore a closed convex cone. Also we already saw that $\mathrm{int}(\mathbb{S}^n_+) = \mathbb{S}^n_{++}$ is non-empty (indeed, $I_n \in \mathbb{S}^n_{++}$).

Next we show that $\mathbb{S}^n_+$ contains no lines. Assume, for a contradiction, that it does contain a line, say $\{X + \alpha D;\ \alpha \in \mathbb{R}\} \subseteq \mathbb{S}^n_+$ for some $X \in \mathbb{S}^n_+$ and some $D \in \mathbb{S}^n \setminus \{0\}$. Then for all $S \in \mathbb{S}^n_+$ and all $\alpha \in \mathbb{R}$ we have

$$0 \leq \langle X + \alpha D, S\rangle = \langle X, S\rangle + \alpha \langle D, S\rangle,$$

and hence $\langle D, S\rangle = 0$ (otherwise we could pick an α to violate the above equation). Then $\langle D, S\rangle = 0 \geq 0$ for all $S \in \mathbb{S}^n_+$, so by the characterization of positive semi-definite matrices we get $D \in \mathbb{S}^n_+$. Similarly, $-D \in \mathbb{S}^n_+$ (because $\{X+\alpha D;\ \alpha \in \mathbb{R}\} = \{X+\alpha(-D);\ \alpha \in \mathbb{R}\}$). So $\lambda(D) \geq 0$ and $\lambda(-D) = -\lambda(D) \geq 0$. Therefore $\lambda(D) = 0$. But for a symmetric matrix, this means $D = 0$ (briefly, $D = Q\,\mathrm{Diag}(\lambda(D))Q^{\mathsf{T}} = Q0Q^{\mathsf{T}} = 0$). Contradiction. Thus $\mathbb{S}^n_+$ is pointed.

So we have that $\mathbb{S}^n_+$ is a pointed, closed cone with non-empty interior. We can now show that it is symmetric. In fact, we already see in equation (3) that $\mathbb{S}^n_+$ is self-dual under the standard trace inner product.

To conclude that $\mathbb{S}^n_+$ is symmetric, it remains to show that it is homogeneous. We will show that for every $X \in \mathrm{int}(\mathbb{S}^n_+) = \mathbb{S}^n_{++}$ that there is a $T_X \in \mathrm{Aut}(\mathbb{S}^n_+)$ such that $T_X(X) = I$. From here, for any pair $X, V \in \mathbb{S}^n_{++}$, $T_X^{-1} \circ T_V \in \mathrm{Aut}(\mathbb{S}^n_{++})$ satisfies $(T_X^{-1} \circ T_V)(V) = T_X^{-1}(T_V(V)) = T_X^{-1}(I) = X$, showing that $\mathbb{S}^n_{++}$ is homogeneous. Let $X \in \mathbb{S}^n_{++}$ and define $T_X : \mathbb{S}^n \to \mathbb{S}^n$ by $T_X(V) = X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$ for all $V \in \mathbb{S}^n$. First, T_X is invertible because its inverse is easily seen to be $T_{X^{-1}}$ (X is invertible because it is positive definite), and it is clearly linear. It is also easy to see that

$$T_X(X) = X^{-\frac{1}{2}} X X^{-\frac{1}{2}} = \left(X^{\frac{1}{2}}\right)^{-1} X^{\frac{1}{2}} X^{\frac{1}{2}} \left(X^{\frac{1}{2}}\right)^{-1} = II = I.$$

Finally, if $V \in \mathbb{S}^n_+$ then for any $h \in \mathbb{R}^n$ we have

$$h^{\mathsf{T}} T_X(V) h = h^{\mathsf{T}} X^{-\frac{1}{2}} V X^{-\frac{1}{2}} h = \left(X^{-\frac{1}{2}} h\right)^{\mathsf{T}} X \left(X^{-\frac{1}{2}} h\right) \geq 0$$

because $X \in \mathbb{S}^n_+$. Hence $T_X(\mathbb{S}^n_+) \subseteq \mathbb{S}^n_+$. To see the reverse inclusion, let $V \in \mathbb{S}^n_+$ and then note

$$T_X(X^{\frac{1}{2}} V X^{\frac{1}{2}}) = X^{-\frac{1}{2}} X^{\frac{1}{2}} V X^{\frac{1}{2}} X^{-\frac{1}{2}} = IVI = V.$$

Thus $T(\mathbb{S}^n_+) = \mathbb{S}^n_+$ and so $T \in \mathrm{Aut}(K)$. This shows that $\mathbb{S}^n_+$ is homogeneous and hence symmetric.

It remains to show $\mathrm{ext}(\mathbb{S}^n_+) = \{hh^{\mathsf{T}};\ h \in \mathbb{R}^n,\ \|h\|_2 = 1\}$. Exercise. ∎

Definition 64. Let $d_1, d_2 \in \mathbb{N}$ and let $K_1 \subseteq \mathbb{R}^{d_1}$, $K_2 \subseteq \mathbb{R}^{d_2}$. We define

$$K_1 \oplus K_2 := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^{d_1+d_2};\ u \in K_1,\ v \in K_2 \right\}.$$

Proposition 65. *Let* $d_1, d_2 \in \mathbb{N}$ *and let* $K_1 \subseteq \mathbb{R}^{d_1}$, $K_2 \subseteq \mathbb{R}^{d_2}$. *Then* $(K_1 \oplus K_2)^* = K_1^* \oplus K_2^*$.

Proof. Let $(x, y) \in \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2}$. First suppose that $(x, y) \in (K_1 \oplus K_2)^*$. We will show that $(x, y) \in K_1^* \oplus K_2^*$. That is, we need to show $x \in K_1^*$ and $y \in K_2^*$. Suppose $z \in K_1$. Then $(z, w) \in K_1 \oplus K_2$ for some $w \in K_2$, so

$$\langle x, z \rangle = \langle (x, y), (z, w) \rangle - \langle y, w \rangle \geq -\langle y, w \rangle \ldots$$

∎

REMARK **66.** Let $n, m \in \mathbb{N}$, $C \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^m$, and $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$. Suppose we have $n_1, \dots, n_r \in \mathbb{N}$ such that $n_1 + \cdots + n_r = n$. Consider the optimization problem (P):

$$\begin{aligned} \inf \quad & \langle C, X\rangle \\ \mathcal{A}(X) &= b \\ X &\in \mathbb{S}_+^{n_1} \oplus \cdots \oplus \mathbb{S}_+^{n_r} \end{aligned}$$

When $n_1 = \cdots = n_r = 1$, we have a regular linear program (since $\mathbb{S}_+^1$ is just $\mathbb{R}_+$). Therefore linear programming (from CO 250) is a special case of what we are doing (CO 471).

Note that the dual is (D):

$$\begin{aligned} \sup \quad & b^T y \\ \mathcal{A}^*(y) + S &= C \\ S &\in \mathbb{S}_+^{n_1} \oplus \cdots \oplus \mathbb{S}_+^{n_r} \end{aligned}$$

This is because $\mathbb{S}_+^{n_1} \oplus \cdots \oplus \mathbb{S}_+^{n_r}$ is self-dual.

THEOREM **67.** [THEOREM 1.17: WEAK DUALITY] *Let $n, m \in \mathbb{N}$, $C \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^m$, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$, and $n_1, \dots, n_r \in \mathbb{N}$ satisfy $n_1 + \cdots + n_r = n$. If $\bar{X}$ and $(\bar{y}, \bar{S})$ are feasible solutions for (P) and for (D) (given above) respectively, then we have*

$$\left\langle C, \bar{X}\right\rangle - \langle b, \bar{y}\rangle = \left\langle \bar{X}, \bar{S}\right\rangle \geq 0.$$

Proof. We simply compute

$$\begin{aligned} \left\langle C, \bar{X}\right\rangle - \langle b, \bar{y}\rangle &= \left\langle \mathcal{A}^*(\bar{y}) + \bar{S}, \bar{X}\right\rangle - \left\langle \mathcal{A}(\bar{X}), \bar{y}\right\rangle \\ &= \left\langle \mathcal{A}^*(\bar{y}), \bar{X}\right\rangle + \left\langle \bar{S}, \bar{X}\right\rangle - \left\langle \mathcal{A}(\bar{X}), \bar{y}\right\rangle \\ &= \left\langle \bar{y}, \mathcal{A}(\bar{X})\right\rangle + \left\langle \bar{S}, \bar{X}\right\rangle - \left\langle \mathcal{A}(\bar{X}), \bar{y}\right\rangle \\ &= \left\langle \bar{S}, \bar{X}\right\rangle = \left\langle \bar{X}, \bar{S}\right\rangle \geq 0. \end{aligned}$$

∎

DEFINITION **68.** Let (P) be an SDP problem, and let $\bar{X}$ be a feasible solution of (P). We say $\bar{X}$ is a *Slater point* of (P) if $X \succ 0$.

PROPOSITION **69.** *Let $n \in \mathbb{N}$ and let $X, S \in \mathbb{S}_+^n$. Then $\langle X, S\rangle = 0$ if and only if $XS = 0$.*

Proof. Suppose first that $XS = 0$. Then $\langle X, S\rangle = \operatorname{Tr}(X^T S) = \operatorname{Tr}(XS) = \operatorname{Tr}(0) = 0$.

Conversely, suppose that $\langle X, S\rangle = 0$. Then

$$\left\langle X^{\frac12}S^{\frac12}, X^{\frac12}S^{\frac12}\right\rangle = \operatorname{Tr}\left(\left(X^{\frac12}S^{\frac12}\right)^T\left(X^{\frac12}S^{\frac12}\right)\right) = \operatorname{Tr}\left(S^{\frac12}XS^{\frac12}\right) = \operatorname{Tr}\left(XS^{\frac12}S^{\frac12}\right) = \operatorname{Tr}\left(X^T S\right) = \langle X, S\rangle = 0,$$

which shows $X^{\frac12}S^{\frac12} = 0$. Thus $XS = X^{\frac12}X^{\frac12}S^{\frac12}S^{\frac12} = X^{\frac12}0S^{\frac12} = 0$. Also note $SX = S^T X^T = (XS)^T = 0^T = 0$. ∎

THEOREM 70. *Let* $n \in \mathbb{N}$, *let* $k \in [n]$, *and let* $X \in \mathbb{S}^n$. *Then*

$$\lambda_k(X) = \min_{\dim(L)=n-k+1} \left\{ \max_{h \in L \setminus \{0\}} \left\{ \frac{h^T X h}{h^T h} \right\} \right\}$$

and

$$\lambda_k(X) = \max_{\dim(L)=k} \left\{ \min_{h \in L \setminus \{0\}} \left\{ \frac{h^T X h}{h^T h} \right\} \right\}.$$

Proof. Exercise. ∎

EXAMPLE 71. Taking $k = 1$ in the above and using either equation, we get that the largest eigenvalue of X is

$$\lambda_1(X) = \max_{h \in \mathbb{R}^n \setminus \{0\}} \left\{ \frac{h^T X h}{h^T h} \right\}.$$

LEMMA 72. [LEMMA 1.22: SCHUR COMPLEMENT] *Let* $k, m \in \mathbb{N}$, *let* $T \in \mathbb{S}^k_{++}$, *let* $X \in \mathbb{S}^m$, *let* $U \in \mathbb{R}^{m \times k}$, *and define*

$$M := \begin{bmatrix} T & U^T \\ U & X \end{bmatrix}.$$

Then $M \succcurlyeq 0$ *if and only if* $X - UT^{-1}U^T \succcurlyeq 0$, *and* $M \succ 0$ *if and only if* $X - UT^{-1}U^T \succ 0$.

Proof. Let

$$L := \begin{bmatrix} I & 0 \\ UT^{-1} & I \end{bmatrix}.$$

It is easy to compute that

$$M = L \begin{bmatrix} T & 0 \\ 0 & X - UT^{-1}U^T \end{bmatrix} L^T.$$

Note that L is invertible (indeed, $\det(L) = \det(I)\det(I) = 1 \neq 0$), so we also have

$$\begin{bmatrix} T & 0 \\ 0 & X - UT^{-1}U^T \end{bmatrix} = L^{-1} M \left(L^{-1}\right)^T.$$

But we already know that if V is a positive (semi-)definite matrix then AVA^T is a positive (semi-)definite matrix (assuming A is invertible in the positive definite case). Therefore the desired results follow by four applications of this observation, as well as the fact that

$$\begin{bmatrix} T & 0 \\ 0 & X - UT^{-1}U^T \end{bmatrix}$$

is positive (semi-)definite if and only if $X - UT^{-1}U^T$ is positive (semi-)definite (because $T \in \mathbb{S}^n_{++}$). ∎

DEFINITION 73. Let $k, m \in \mathbb{N}$, let $T \in \mathbb{S}^k_{++}$, let $X \in \mathbb{S}^m$ and let $U \in \mathbb{R}^{m\times k}$. Then $X - UT^{-1}U^T$ is called the *Schur Complement* of T in

$$\begin{bmatrix} T & U^T \\ U & X \end{bmatrix}.$$

REMARK 74. In the special case that $T = \begin{bmatrix} 1 \end{bmatrix}$, we see

$$\begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0 \quad \text{if and only if} \quad X - xx^T \succcurlyeq 0.$$

DEFINITION 75. Let $n, m, p, q \in \mathbb{N}$, let $U \in \mathbb{R}^{m\times n}$, and let $V \in \mathbb{R}^{p\times q}$. The *Kronecker Product* (or *Tensor Product*) of U and V is

$$U \otimes V := \begin{bmatrix} U_{1,1}V & \cdots & U_{1,n}V \\ \vdots & \ddots & \vdots \\ U_{m,1}V & \cdots & U_{m,n}V \end{bmatrix} \in \mathbb{R}^{mp\times nq}.$$

PROPOSITION 76. *Let $n, m \in \mathbb{N}$, let $U \in \mathbb{S}^n$ and let $V \in \mathbb{S}^m$. If $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_m$ are the eigenvalues of U and of V respectively, then $\lambda_1\mu_1, \ldots, \lambda_1\mu_m, \ldots, \lambda_n\mu_1, \ldots, \lambda_n\mu_m$ are the eigenvalues of $U \otimes V$ (not necessarily in decreasing order). Moreover, if $u^{(1)}, \ldots, u^{(n)}$ and $v^{(1)}, \ldots, v^{(m)}$ are the eigenvectors of U and V respectively, then $u^{(1)} \otimes v^{(1)}, \ldots, u^{(1)} \otimes v^{(m)}, \ldots, u^{(n)} \otimes v^{(1)}, \ldots, u^{(n)} \otimes v^{(m)}$ are the eigenvectors of $U \otimes V$.*

Proof. Let $(i, j) \in [n]\times[m]$. Clearly $u^{(i)} \otimes v^{(j)}$ is non-zero because $u^{(i)}$ and $v^{(j)}$ are non-zero (as they are eigenvectors). We will show that $(U \otimes V)(u^{(i)} \otimes v^{(j)}) = \lambda_i\mu_j(u^{(i)} \otimes v^{(j)})$.

So we calculate

$$\begin{aligned}
(U \otimes V)\left(u^{(i)} \otimes v^{(j)}\right) &= \begin{bmatrix} U_{1,1}V & \cdots & U_{1,n}V \\ \vdots & \ddots & \vdots \\ U_{m,1}V & \cdots & U_{m,n}V \end{bmatrix} \begin{bmatrix} u^{(i)}_1 v^{(j)} \\ \vdots \\ u^{(i)}_n v^{(j)} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^n (U_{1,k}V)\left(u^{(i)}_k v^{(j)}\right) \\ \vdots \\ \sum_{k=1}^n (U_{n,k}V)\left(u^{(i)}_k v^{(j)}\right) \end{bmatrix} \\
&= \begin{bmatrix} (Vv^{(j)}) \sum_{k=1}^n U_{1,k}u^{(i)}_k \\ \vdots \\ (Vv^{(j)}) \sum_{k=1}^n U_{n,k}u^{(i)}_k \end{bmatrix} \\
&= \begin{bmatrix} \mu_j v^{(j)} \lambda_i u^{(i)}_1 \\ \vdots \\ \mu_j v^{(j)} \lambda_i u^{(i)}_n \end{bmatrix}
\end{aligned}$$

$$= \lambda_i \mu_j \begin{bmatrix} u_1^{(i)} v^{(j)} \\ \vdots \\ u_n^{(i)} v^{(j)} \end{bmatrix}$$
$$= \lambda_i \mu_j \left(u^{(i)} \otimes v^{(j)} \right).$$

∎

Corollary 77. *Let* $n, m \in \mathbb{N}$*, let* $U \in \mathbb{S}^n$*, and let* $V \in \mathbb{S}^m$*. If* $U, V \succeq 0$ *then* $U \otimes V \succeq 0$*, and if* $U, V \succ 0$ *then* $U \otimes V \succeq 0$*.*

Proof. First note that it is easy to check that $U \otimes V \in \mathbb{S}^{nm}$.

Suppose $U \in \mathbb{S}^n_+$ and $V \in \mathbb{S}^m_+$. Then all the eigenvalues of U and of V are non-negative. By the previous proposition, all the eigenvalues of $U \otimes V$ are non-negative. Thus $U \otimes V \in \mathbb{S}^{nm}_+$.

Suppose $U \in \mathbb{S}^n_{++}$ and $V \in \mathbb{S}^m_{++}$. Then all the eigenvalues of U and of V are positive. By the previous proposition, all the eigenvalues of $U \otimes V$ are positive. Thus $U \otimes V \mathbb{S}^{nm}_{++}$. ∎

Definition 78. Let $n \in \mathbb{N}$ and let $X, S \in \mathbb{R}^{n \times n}$. The *Hadamard Product* of X and S is defined by

$$(X \odot S)_{i,j} := X_{i,j} S_{i,j}$$

for all $(i, j) \in [n]^2$.

Corollary 79. *Let* $n \in \mathbb{N}$ *and let* $X, S \in \mathbb{S}^n$*. If* $X, S \succeq 0$ *then* $X \odot S \succeq 0$*, and if* $X, S \succ 0$ *then* $X \odot S \succ 0$*.*

Proof. Suppose $X, S \in \mathbb{S}^n_+$. We saw that $X \otimes S \in \mathbb{S}^n_+$. Moreover, $X \odot S$ is a symmetric minor of $X \otimes S$ (check) and therefore $X \odot S \in \mathbb{S}^n_+$ by lemma (40).

Suppose $X, S \in \mathbb{S}^n_{++}$. We saw that $X \otimes S \in \mathbb{S}^n_{++}$. Again, $X \odot S$ is a symmetric minor of $X \otimes S$, and therefore $X \odot S \in \mathbb{S}^n_{++}$ by an unstated lemma which is analogous to lemma (40). ∎

Definition 80. Let $n \in \mathbb{N}$. Define $\mathrm{vec} : \mathbb{R}^{n \times n} \to \mathbb{R}^{n^2}$ by

$$\mathrm{vec}(X) := \begin{bmatrix} X_{1,1} \\ \vdots \\ X_{n,1} \\ \vdots \\ X_{1,n} \\ \vdots \\ X_{n,n} \end{bmatrix}$$

and $\text{s2vec} : \mathbb{S}^n \to \mathbb{R}^{\frac{n(n-1)}{2}}$ by

$$\text{s2vec}(X) := \begin{bmatrix} X_{1,1} \\ \sqrt{2}X_{2,1} \\ \vdots \\ \sqrt{2}X_{n,1} \\ X_{2,2} \\ \sqrt{2}X_{3,2} \\ \vdots \\ \sqrt{2}X_{n,2} \\ \vdots \\ X_{n,n} \end{bmatrix}.$$

2 Duality Theory

THEOREM 81. [THEOREM 2.8: A SEPARATION THEOREM] *Let* $n \in \mathbb{N}$ *and let* $G \subseteq \mathbb{R}^n$*. If* G *is a non-empty, closed convex set with* $0 \notin G$*, then there exists an* $a \in \mathbb{R}^n \setminus \{0\}$ *and an* $\alpha \in \mathbb{R}_{++}$ *such that* $a^T x \geq \alpha$ *for all* $x \in G$*.*

Proof. Since G is non-empty, let $\bar{x} \in G$. Set

$$G_{\bar{x}} := \{x \in G; \|x\|_2 \leq \|\bar{x}\|_2\}.$$

Note that $G_{\bar{x}} = G \cap \{x \in \mathbb{R}^n; \|x\|_2 \leq \|\bar{x}\|_2\}$ in the intersection of compact convex sets and is hence compact and convex. Moreover, it is non-empty because $\bar{x} \in G_{\bar{x}}$, and $0 \notin G_{\bar{x}}$ since $0 \notin G$. As such, there exists a unique point a in $G_{\bar{x}}$ which is closest to the origin. (Indeed, consider minimizing the continuous, strictly convex function $\|\cdot\|_2^2$ on the non-empty compact convex set $G_{\bar{x}}$.)

Since $a \in G_{\bar{x}}$ and $0 \notin G_{\bar{x}}$, $a \neq\in \mathbb{R}^n \setminus \{0\}$. Set $\alpha := \|a\|_2^2 > 0$. Let $x \in G$. Since G is convex, for every $\lambda \in (0,1)$ we have $\lambda x + (1-\lambda)a \in G$. By choice of a, we have $\|a\|_2 \leq \|z\|_2^2$ for all $z \in G$. Therefore

$$\|\lambda x + (1-\lambda)a\|_2^2 \geq \|a\|_2^2.$$

But expanding the left side yields

$$\|\lambda(x-a) + a\|_2^2 = \lambda^2 \|x-a\|_2^2 + 2\lambda \langle x-a, a\rangle + \|a\|_2^2.$$

Hence the first inequality becomes

$$\lambda^2 \|x-a\|_2^2 + \lambda \langle x-a, a\rangle \geq 0,$$

which shows, since $\lambda > 0$,

$$(x-a)^T a \geq -\tfrac{\lambda}{2} \|x-a\|_2^2.$$