2.8

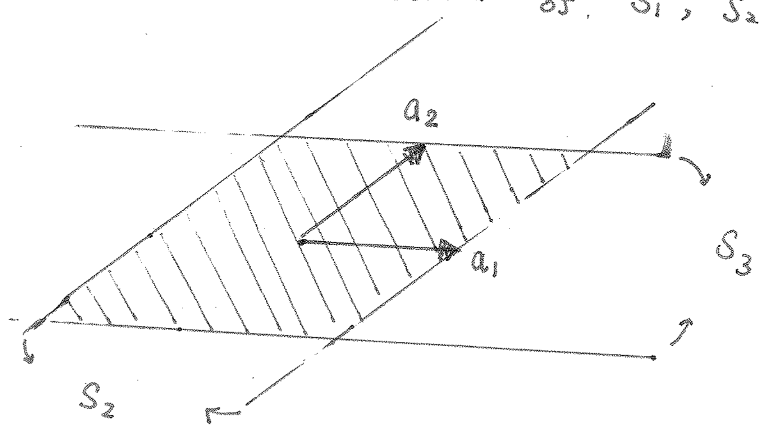
(a) $S = \{ y_1 a_1 + y_2 a_2 \mid -1 \le y_1 \le 1,\ -1 \le y_2 \le 1 \}$, $a_1, a_2 \in \mathbb{R}^n$ is a polyhedron.

First, let $S_1 = \{ y_1 a_1 + y_2 a_2 \mid y_1, y_2 \in \mathbb{R},\ a_1, a_2 \in \mathbb{R}^n \}$ be the plane defined by a_1, a_2.

Second, let $S_2 = \{ z + y_1 a_1 + y_2 a_2 \mid -1 \le y_1 \le 1,\ z, a_1, a_2 \in \mathbb{R}^n,\ z^T a_1 = z^T a_2 = 0 \}$ be the slab parallel to a_2 and orthogonal to S_1.

Third, let $S_3 = \{ z + y_1 a_1 + y_2 a_2 \mid -1 \le y_2 \le 1,\ z, a_1, a_2 \in \mathbb{R}^n,\ z^T a_1 = z^T a_2 = 0 \}$ be the slab parallel to a_1 and orthogonal to S_1.

Then the S is the intersection of S_1, S_2, S_3, as the below figure.



Then we can describe S_1, S_2, S_3 with linear inequalities as follow:

S_1: $u_i^T x = 0$ for $i = 1, \dots, n-2$.

u_i are $n-2$ vectors orthogonal to a_1 and a_2.

S_2: let b_1 be a vector in S_1 and orthogonal to a_2, and we can set b_1 as

$b_1 = a_1 - \dfrac{a_1^T a_2}{\|a_2\|_2^2} a_2$. Then $x \in S_2$ if and only if $-|b_1^T a_1| \le b_1^T x \le |b_1^T a_1|$

S_3: let b_2 be a vector in S_1 and orthogonal to a_1, and we can set b_2 as

$b_2 = a_2 - \dfrac{a_2^T a_1}{\|a_1\|_2^2} a_1$. Then $x \in S_3$ if and only if $-|b_2^T a_2| \le b_2^T x \le |b_2^T a_2|$

Put them together, we can denote S as the following linear inequalities:

$$u_i^T x = 0 \quad , \quad i = 1, \dots, n-2$$

$$b_1^T x \le |b_1^T a_1|$$

$$-b_1^T x \le |b_1^T a_1|$$

$$b_2^T x \le |b_2^T a_2|$$

$$-b_2^T x \le |b_2^T a_2|$$

(b) We can set $A_1 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, $A_2 = \begin{bmatrix} a_1^2 \\ a_2^2 \\ \vdots \\ a_n^2 \end{bmatrix}$, then we can describe S as

$S = \{ x \in \mathbb{R}^n \mid x \ge 0 , \; I^T x = 1 , \; A_1^T x = b_1 , \; A_2^T x = b_2 \}$ where

$a_1, \dots, a_n \in \mathbb{R}$ and $b_1, b_2 \in \mathbb{R}$. By definition, polyhedron is intersection of finite number of halfspaces and hyperplanes, so S is a polyhedron.

(c) By Cauchy - Schwarz inequality, we know that $x^T y \le \|x\|_2 \cdot \|y\|_2$. And by definition of S, we know that $x^T y \le 1$ for all $\|y\|_2 = 1$. Hence, we have $\|x\|_2 \le 1$.

So S is the intersection of the unit ball $\{ x \mid \|x\|_2 \le 1 \}$ and the nonnegative orthant $\mathbb{R}^n_+$. By the definition of polyhedron, we have that S is not a polyhedron.

(d) First of all, we have to prove $x^T y \le 1$ for all y with $\sum_{i=1}^n |y_i| = 1 \iff |x_i| \le 1 , \; i = 1, \dots, n$

Suppose that $|x_i| \le 1$ for all i. Then $x^T y = \sum_i x_i y_i \le \sum_i |x_i| |y_i| \le \sum_i |y_i| = 1$ if $\sum_i |y_i| = 1$.

Conversely, suppose that x is a nonzero vector that satisfies $x^T y \le 1$ for all y with $\sum_i |y_i| = 1$.

We can make y : let j be an index for which $|x_j| = \max_i |x_i|$ and take $y_j = 1$ if $x_j > 0$, $y_j = -1$ if $x_j < 0$, and $y_i = 0$ for $i \ne j$.

With the choice of y, we have $x^T y = \sum_i x_i y_i = y_j x_j = |x_j| = \max_i |x_i|$

Therefore, we must have $\max_i |x_i| \le 1$.

All this implies that we can describe S by a finite number of linear inequalities.

The set S is the intersection of the nonnegative orthant with the set $\{ x \mid -1 \le x \le 1 \}$, which is the solution of the $2n$ linear inequalities:

$$-x_i \le 0 , \quad i = 1, \dots, n$$

$$x_i \le 1 , \quad i = 1, \dots, n$$

Therefore, S is a polyhedron.

2.13

We have $XX^T \succeq 0$ and rank $(XX^T) = k$. A positive combination of such matrices can have rank up to n, but never less than k. Let A, B be positive semi definite matrices of rank k. Suppose $v \in Null(A+B)$, then we have

$$(A+B)v = 0 \Leftrightarrow v^T(A+B)v = 0 \Leftrightarrow v^TAv + v^TBv = 0$$

Since A, B are positive semi definite and $v^TAv + v^TBv = 0$, we have

$$v^TAv = 0 \Leftrightarrow Av = 0 \quad , \quad v^TBv = 0 \Leftrightarrow Bv = 0.$$

Hence, any vector in $Null(A+B)$ must be in $Null(A)$ and $Null(B)$, which means that $\dim(Null(A+B))$ cannot be greater than $\dim(Null(A))$ and $\dim(Null(B))$.

Therefore, we conclude that $rank(A+B) \geq k$ for any A, B such that $rank(A, B) = k$ and $A, B \succeq 0$.

It follows that the conic hull of the set of rank-k outer products is the set of positive semidefinite matrices of rank greater than or equal to k, along with the zero matrix.

2.22: Following the hint, we have to prove that $S = \{x - y \mid x \in C, y \in D\}$ is convex first.

Assume that $e, f \in S$, $a, b \in C$, $c, d \in D$ such that $e = a - c$, $f = b - d$.

Then we have $t \cdot e + (1-t)f = t(a-c) + (1-t)(b-d)$

$$= \big(ta + (1-t)b\big) - \big(tc + (1-t)d\big) \in C - D = S$$

Therefore, we have confirmed that S is convex.

Since C and D are disjoint, $0 \notin S$. First, suppose $0 \notin$ Closure of S ($\mathbf{cl}\, S$).

The partial seperating hyperplane in 2.5.1 applies to the sets $\{0\}$ and $\mathbf{cl}\, S$, so there exists an $a \neq 0$ such that $a^T(x-y) > 0$ for all $x - y \in \mathbf{cl}\, S$. And this also holds for all $x - y \in S$, i.e. $a^Tx > a^Ty$ for all $x \in C$ and $y \in D$.

Next, assume $0 \in \mathbf{cl}\, S$. Since $0 \notin S$, 0 has to be in the boundary of S.

If S has empty interior, it is contained in a hyperplane $\{z \mid a^Tz = b\}$, which must include the origin, hence $b = 0$. That is $a^Tx = a^Ty$ for all $x \in C$ and $y \in D$

So we have a trivial seperating hyperplane.

If S has nonempty interior, we consider the set $S_{-\epsilon} = \{z \mid B(z,\epsilon) \subseteq S\}$ where $B(z,\epsilon)$ is the Euclidean ball with center z and radius $\epsilon > 0$. $S_{-\epsilon}$ is S shrunk by ϵ, so $\mathbf{cl}\, S_{-\epsilon}$ is closed and convex and does not contain 0. By partial seperating hyperplane result, it is strictly seperated from $\{0\}$ by at least one hyperplane with normal vector $a(\epsilon)$:

$a(\epsilon)^T z > 0$ for all $z \in S_{-\epsilon}$ and we can assume $\|a(\epsilon)\|_2 = 1$.

Now let ϵ_k, $k=1,2,\dots$ be a sequence of positive values of ϵ_k with $\lim_{k\to\infty} \epsilon_k = 0$. Since $\|a(\epsilon_k)\|_2 = 1$ for all k, the sequence $a(\epsilon_k)$ contains a convergent subsequence, and we can denote its limit by $\bar{a}$. We have $a(\epsilon_k)^T z > 0$ for all $z \in S_{-\epsilon_k}$ for all k. Therefore, $\bar{a}^T z > 0$ for all $z \in \mathbf{int}\, S$ and $\bar{a}^T z \geq 0$ for all $z \in S$, i.e.,

$\bar{a}^T x \geq \bar{a}^T y$ for all $x \in C$ and $y \in D$.

A1.5

(a) Since C and D are convex, and the intersection operation preserve convexity. Therefore, $C \cap D$ is convex. To prove $C \cap D$ is a cone, suppose $x \in C \cap D$. It implies that $x \in C$ and $x \in D$, which also implies $\theta x \in C$ and $\theta x \in D$ for any $\theta \geq 0$ since C, D are cones. Thus, $\theta x \in C \cap D$ for any $\theta \geq 0$. Hence, $C \cap D$ is a convex cone.

From the textbook, we know that a dual cone is always convex, so C^* and D^* are convex. And $C^* + D^*$ is the convex hull of $C^* \cup D^*$, which is a convex cone.

(b) Suppose $x \in C^* + D^*$, then we can denote x as $x = u + v$, where $u \in C^*$ and $v \in D^*$. By definition of dual cone, we know that $u^T y \geq 0$ for all $y \in C$ and $v^T y \geq 0$ for all $y \in D$. It implies that $x^T y = (u^T + v^T) y = u^T y + v^T y \geq 0$ for all $y \in C \cap D$. Therefore, $x \in (C \cap D)^*$, and so $(C \cap D)^* \supseteq C^* + D^*$.

(c) From (a), we know that $C \cap D$ and $C^* + D^*$ are convex cones. Hence, we have $(C \cap D)^{**} = C \cap D$ and $(C^* + D^*)^{**} = C^* + D^*$. Suppose $x \in (C^* + D^*)^*$. It implies that $x^T y \geq 0$ for all $y \in C^* + D^*$ and we can write y as $y = u + v$, $u \in C^*$ and $v \in D^*$. This can be rewritten as $x^T y = x^T (u + v) = x^T u + x^T v \geq 0$ for all $u \in C^*$ and $v \in D^*$. Since $0 \in C^*$ and $0 \in D^*$, we can set $v = 0$ such that $x^T u \geq 0$ for all $u \in C^*$, and set $u = 0$ such that $x^T v \geq 0$ for all $v \in D^*$. This implies that $x \in C^{**} = C$ and $x \in D^{**} = D$, so $x \in C \cap D$. Hence, $C \cap D \supseteq (C^* + D^*)^*$ which implies $(C \cap D)^* \subseteq (C^* + D^*)^{**} = C^* + D^*$. We have showed $(C \cap D)^* \supseteq C^* + D^*$ and

$(C \cap D)^* \subseteq C^* + D^*$, which implies $(C \cap D)^* = C^* + D^*$.

(d)

$V = \{x \mid Ax \geq 0\}$ where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$. V can be written as:

$$V = \{x \mid Ax \geq 0\} = \{x \mid a_1 x \geq 0\} \cap \{x \mid a_2 x \geq 0\} \cap \cdots \cap \{x \mid a_m x \geq 0\},$$

$a_1, \dots, a_m \in \mathbb{R}^{1 \times n}$.

Using the previous results, we can write V^* as:

$$V^* = \{x \mid a_1 x \geq 0\}^* + \cdots + \{x \mid a_m x \geq 0\}^*$$

The dual of $\{x \mid a_i x \geq 0\}$ is:

$$V_i^* = \{x \mid a_i x \geq 0\}^* = \{y \mid y^T x \geq 0 \text{ for all } x \in V_i\} = \{v a_i^T \mid (v a_i^T)^T x \geq 0 \text{ for all } x \in V_i\}$$

$$= \{v a_i^T \mid v a_i x \geq 0 \text{ for all } x \in V_i\} = \{v \cdot a_i^T \mid v \geq 0\}$$

Therefore,

$$V^* = \{v a_1^T \mid v \geq 0\} + \cdots + \{v a_m^T \mid v \geq 0\} = \{v_1 a_1^T + \cdots + v_m a_m^T \mid v_i \geq 0 \text{ for } i = 1, \dots, m\}$$

$$= \{A^T v \mid v \geq 0\}$$

A.1.9

(a)

Suppose $C_1 \in C^n$, $C_2 \in C^n$ and $0 \le \theta \le 1$. Then let $C = \theta C_1 + (1-\theta) C_2$.

$x^T C x = x^T \left(\theta C_1 + (1-\theta) C_2\right) x = \theta \cdot x^T C_1 x + (1-\theta) x^T C_2 x$

Since $C_1 \in C^n$, $C_2 \in C^n$, we have $x^T C_1 x \ge 0$, $x^T C_2 x \ge 0$. And since $0 \le \theta \le 1$, we have $\theta x^T C_1 x \ge 0$ and $(1-\theta) x^T C_2 x \ge 0$, which implies $x^T C x \ge 0$ means that $C \in S_+^n$.

And since $C_1 \in C^n$ and $C_2 \in C^n$, we have $C_{1,ii} = 1$ for $i = 1, \dots, n$ and $C_{2,ii} = 1$ for $i = 1, \dots, n$.

Then $C_{ii} = \theta C_{1,ii} + (1-\theta) C_{2,ii} = \theta \times 1 + (1-\theta) \times 1 = 1$ for $i = 1, \dots, n$.

Hence, $C \in C^n$ which means that C^n is a convex set.

(b)

Suppose C_1 and C_2 are nonnegative correlation matrices and $0 \le \theta \le 1$. Then let $C = \theta C_1 + (1-\theta) C_2$.

Since $C_1 \in C^n$, $C_2 \in C^n$ where C^n is a convex set, $C = \theta C_1 + (1-\theta) C_2 \in C^n$ for any $0 \le \theta \le 1$.

And C_1, $C_2 \in$ nonnegative correlation matrices, we have $C_{1,ij} \ge 0$ and $C_{2,ij} \ge 0$ for $i,j = 1, \dots, n$.

Then we have $C_{ij} = \theta C_{1,ij} + (1-\theta) C_{2,ij} \ge 0$ for $i,j = 1, \dots, n$.

Hence, C is also a nonnegative correlation matrix, which implies that $\{C \in C^n \mid C_{ij} \ge 0, i,j = 1, \dots, n\}$ is a convex set.

(c)

Suppose C_1 and C_2 are highly correlated correlation matrices and $0 \le \theta \le 1$. Then set $C = \theta C_1 + (1-\theta) C_2$

Since C^n is a convex set and $C_1 \in C^n$, $C_2 \in C^n$, $C = \theta C_1 + (1-\theta) C_2 \in C^n$ for any $0 \le \theta \le 1$.

And $C_{1,ij} \ge 0.8$ for $i,j = 1, \dots, n$, $C_{2,ij} \ge 0.8$ for $i,j = 1, \dots, n$. This implies that

$C_{ij} = \theta C_{1,ij} + (1-\theta) C_{2,ij} \ge \theta \times 0.8 + (1-\theta) \times 0.8 = 0.8$ for $i,j = 1, \dots, n$.

Hence, C is also a highly correlated correlation matrix, which implies that $\{C \in C^n \mid C_{i,j} \ge 0.8, i,j = 1, \dots, n\}$ is a convex set.