EE364a Homework 1 solutions

2.1 Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k.) *Hint.* Use induction on k.

Solution. This is readily shown by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x_1, x_2, x_3 \in C$, and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_1, \theta_2, \theta_3 \geq 0$. We will show that $y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$. At least one of the θ_i is not equal to one; without loss of generality we can assume that $\theta_1 \neq 1$. Then we can write

$$y = \theta_1 x_1 + (1 - \theta_1)(\mu_2 x_2 + \mu_3 x_3)$$

where $\mu_2 = \theta_2/(1 - \theta_1)$ and $\mu_2 = \theta_3/(1 - \theta_1)$. Note that $\mu_2, \mu_3 \geq 0$ and

$$\mu_1 + \mu_2 = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1.$$

Since C is convex and $x_2, x_3 \in C$, we conclude that $\mu_2 x_2 + \mu_3 x_3 \in C$. Since this point and x_1 are in C, $y \in C$.

2.2 Show that a set is convex if and only if its intersection with any line is convex. Show that a set is affine if and only if its intersection with any line is affine.

Solution. We prove the first part. The intersection of two convex sets is convex. Therefore if S is a convex set, the intersection of S with a line is convex.

Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

2.6 *When does one halfspace contain another?* Give conditions under which

$$\{x \mid a^T x \leq b\} \subseteq \{x \mid \tilde{a}^T x \leq \tilde{b}\}$$

(where $a \neq 0$, $\tilde{a} \neq 0$). Also find the conditions under which the two halfspaces are equal.

Solution. Let $\mathcal{H} = \{x \mid a^T x \leq b\}$ and $\tilde{\mathcal{H}} = \{x \mid \tilde{a}^T x \leq \tilde{b}\}$. The conditions are:

- $\mathcal{H} \subseteq \tilde{\mathcal{H}}$ if and only if there exists a $\lambda > 0$ such that $\tilde{a} = \lambda a$ and $\tilde{b} \geq \lambda b$.
- $\mathcal{H} = \tilde{\mathcal{H}}$ if and only if there exists a $\lambda > 0$ such that $\tilde{a} = \lambda a$ and $\tilde{b} = \lambda b$.

Let us prove the first condition. The condition is clearly sufficient: if $\tilde{a} = \lambda a$ and $\tilde{b} \geq \lambda b$ for some $\lambda > 0$, then

$$a^T x \leq b \implies \lambda a^T x \leq \lambda b \implies \tilde{a}^T x \leq \tilde{b},$$

i.e., $\mathcal{H} \subseteq \tilde{\mathcal{H}}$.

To prove necessity, we distinguish three cases. First suppose a and $\tilde{a}$ are not parallel. This means we can find a v with $a^T v = 0$ and $\tilde{a}^T v \neq 0$. Let $\hat{x}$ be any point in the intersection of $\mathcal{H}$ and $\tilde{\mathcal{H}}$, *i.e.*, $a^T \hat{x} \leq b$ and $\tilde{a}^T x \leq \tilde{b}$. We have $a^T(\hat{x} + tv) = a^T \hat{x} \leq b$ for all $t \in \mathbf{R}$. However $\tilde{a}^T(\hat{x} + tv) = \tilde{a}^T \hat{x} + t\tilde{a}^T v$, and since $\tilde{a}^T v \neq 0$, we will have $\tilde{a}^T(\hat{x} + tv) > \tilde{b}$ for sufficiently large $t > 0$ or sufficiently small $t < 0$. In other words, if a and $\tilde{a}$ are not parallel, we can find a point $\hat{x} + tv \in \mathcal{H}$ that is not in $\tilde{\mathcal{H}}$, *i.e.*, $\mathcal{H} \not\subseteq \tilde{\mathcal{H}}$.

Next suppose a and $\tilde{a}$ are parallel, but point in opposite directions, *i.e.*, $\tilde{a} = \lambda a$ for some $\lambda < 0$. Let $\hat{x}$ be any point in $\mathcal{H}$. Then $\hat{x} - ta \in \mathcal{H}$ for all $t \geq 0$. However for t large enough we will have $\tilde{a}^T(\hat{x} - ta) = \tilde{a}^T \hat{x} + t\lambda\|a\|_2^2 > \tilde{b}$, so $\hat{x} - ta \notin \tilde{\mathcal{H}}$. Again, this shows $\mathcal{H} \not\subseteq \tilde{\mathcal{H}}$.

Finally, we assume $\tilde{a} = \lambda a$ for some $\lambda > 0$ but $\tilde{b} < \lambda b$. Consider any point $\hat{x}$ that satisfies $a^T \hat{x} = b$. Then $\tilde{a}^T \hat{x} = \lambda a^T \hat{x} = \lambda b > \tilde{b}$, so $\hat{x} \notin \tilde{\mathcal{H}}$.

The proof for the second part of the problem is similar.

2.9 *Voronoi sets and polyhedral decomposition.* Let $x_0, \ldots, x_K \in \mathbf{R}^n$. Consider the set of points that are closer (in Euclidean norm) to x_0 than the other x_i, *i.e.*,

$$V = \{x \in \mathbf{R}^n \mid \|x - x_0\|_2 \leq \|x - x_i\|_2,\ i = 1, \ldots, K\}.$$

V is called the *Voronoi region* around x_0 with respect to $x_1, \ldots, x_K$.

(a) Show that V is a polyhedron. Express V in the form $V = \{x \mid Ax \preceq b\}$.

(b) Conversely, given a polyhedron P with nonempty interior, show how to find $x_0, \ldots, x_K$ so that the polyhedron is the Voronoi region of x_0 with respect to $x_1, \ldots, x_K$.

(c) We can also consider the sets

$$V_k = \{x \in \mathbf{R}^n \mid \|x - x_k\|_2 \leq \|x - x_i\|_2,\ i \neq k\}.$$

The set V_k consists of points in $\mathbf{R}^n$ for which the closest point in the set $\{x_0, \ldots, x_K\}$ is x_k.

The sets $V_0, \ldots, V_K$ give a polyhedral decomposition of $\mathbf{R}^n$. More precisely, the sets V_k are polyhedra, $\bigcup_{k=0}^K V_k = \mathbf{R}^n$, and $\mathbf{int}\, V_i \cap \mathbf{int}\, V_j = \emptyset$ for $i \neq j$, *i.e.*, V_i and V_j intersect at most along a boundary.

Suppose that $P_1, \ldots, P_m$ are polyhedra such that $\bigcup_{i=1}^m P_i = \mathbf{R}^n$, and $\mathbf{int}\, P_i \cap \mathbf{int}\, P_j = \emptyset$ for $i \neq j$. Can this polyhedral decomposition of $\mathbf{R}^n$ be described as the Voronoi regions generated by an appropriate set of points?

Solution.

(a) x is closer to x_0 than to x_i if and only if

$$\begin{aligned}
\|x - x_0\|_2 \le \|x - x_i\|_2 &\iff (x - x_0)^T(x - x_0) \le (x - x_i)^T(x - x_i) \\
&\iff x^T x - 2x_0^T x + x_0^T x_0 \le x^T x - 2x_i^T x + x_i^T x_i \\
&\iff 2(x_i - x_0)^T x \le x_i^T x_i - x_0^T x_0,
\end{aligned}$$

which defines a halfspace. We can express V as $V = \{x \mid Ax \preceq b\}$ with

$$A = 2\begin{bmatrix} (x_1 - x_0)^T \\ (x_2 - x_0)^T \\ \vdots \\ (x_K - x_0)^T \end{bmatrix}, \qquad b = \begin{bmatrix} x_1^T x_1 - x_0^T x_0 \\ x_2^T x_2 - x_0^T x_0 \\ \vdots \\ x_K^T x_K - x_0^T x_0 \end{bmatrix}.$$

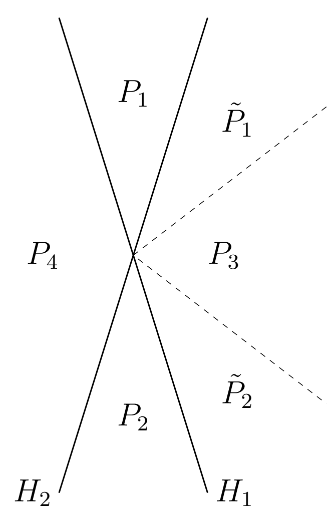
(b) Conversely, suppose $V = \{x \mid Ax \preceq b\}$ with $A \in \mathbf{R}^{K \times n}$ and $b \in \mathbf{R}^K$. We can pick any $x_0 \in \{x \mid Ax \prec b\}$, and then construct K points x_i by taking the mirror image of x_0 with respect to the hyperplanes $\{x \mid a_i^T x = b_i\}$. In other words, we choose x_i of the form $x_i = x_0 + \lambda a_i$, where λ is chosen in such a way that the distance of x_i to the hyperplane defined by $a_i^T x = b_i$ is equal to the distance of x_0 to the hyperplane:

$$b_i - a_i^T x_0 = a_i^T x_i - b_i.$$

Solving for λ, we obtain $\lambda = 2(b_i - a_i^T x_0)/\|a_i\|_2^2$, and

$$x_i = x_0 + \frac{2(b_i - a_i^T x_0)}{\|a_i\|^2} a_i.$$

(c) A polyhedral decomposition of $\mathbf{R}^n$ can not always be described as Voronoi regions generated by a set of points $\{x_1, \ldots, x_m\}$. The figure shows a counterexample in $\mathbf{R}^2$.



$\mathbf{R}^2$ is decomposed into 4 polyhedra $P_1, \dots, P_4$ by 2 hyperplanes H_1, H_2. Suppose we arbitrarily pick $x_1 \in P_1$ and $x_2 \in P_2$. $x_3 \in P_3$ must be the mirror image of x_1 and x_2 with respect to H_2 and H_1, respectively. However, the mirror image of x_1 with respect to H_2 lies in $\tilde{P}_1$, and the mirror image of x_2 with respect to H_1 lies in $\tilde{P}_2$, so it is impossible to find such an x_3.

2.12 Which of the following sets are convex?

(a) A *slab*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha \le a^T x \le \beta\}$.

(b) A *rectangle*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \le x_i \le \beta_i,\ i = 1, \dots, n\}$. A rectangle is sometimes called a *hyperrectangle* when $n > 2$.

(c) A *wedge*, *i.e.*, $\{x \in \mathbf{R}^n \mid a_1^T x \le b_1,\ a_2^T x \le b_2\}$.

(d) The set of points closer to a given point than a given set, *i.e.*,

$$\{x \mid \|x - x_0\|_2 \le \|x - y\|_2 \text{ for all } y \in S\}$$

where $S \subseteq \mathbf{R}^n$.

(e) The set of points closer to one set than another, *i.e.*,

$$\{x \mid \mathbf{dist}(x, S) \le \mathbf{dist}(x, T)\},$$

where $S, T \subseteq \mathbf{R}^n$, and

$$\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}.$$

(f) The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbf{R}^n$ with S_1 convex.

(g) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, *i.e.*, the set $\{x \mid \|x - a\|_2 \le \theta \|x - b\|_2\}$. You can assume $a \ne b$ and $0 \le \theta \le 1$.

Solution.

(a) A slab is an intersection of two halfspaces, hence it is a convex set (and a polyhedron).

(b) As in part (a), a rectangle is a convex set and a polyhedron because it is a finite intersection of halfspaces.

(c) A wedge is an intersection of two halfspaces, so it is convex set. It is also a polyhedron. It is a cone if $b_1 = 0$ and $b_2 = 0$.

(d) This set is convex because it can be expressed as

$$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \le \|x - y\|_2\},$$

i.e., an intersection of halfspaces. (For fixed y, the set

$$\{x \mid \|x - x_0\|_2 \le \|x - y\|_2\}$$

is a halfspace; see exercise 2.9).

(e) In general this set is not convex, as the following example in $\mathbf{R}$ shows. With $S = \{-1, 1\}$ and $T = \{0\}$, we have

$$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\} = \{x \in \mathbf{R} \mid x \leq -1/2 \text{ or } x \geq 1/2\}$$

which clearly is not convex.

(f) This set is convex. $x + S_2 \subseteq S_1$ if $x + y \in S_1$ for all $y \in S_2$. Therefore

$$\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\} = \bigcap_{y \in S_2} (S_1 - y),$$

the intersection of convex sets $S_1 - y$.

(g) The set is convex, in fact a ball.

$$\begin{aligned} &\{x \mid \|x - a\|_2 \leq \theta \|x - b\|_2\} \\ &= \{x \mid \|x - a\|_2^2 \leq \theta^2 \|x - b\|_2^2\} \\ &= \{x \mid (1 - \theta^2) x^T x - 2(a - \theta^2 b)^T x + (a^T a - \theta^2 b^T b) \leq 0\} \end{aligned}$$

If $\theta = 1$, this is a halfspace. If $\theta < 1$, it is a ball

$$\{x \mid (x - x_0)^T (x - x_0) \leq R^2\},$$

with center x_0 and radius R given by

$$x_0 = \frac{a - \theta^2 b}{1 - \theta^2}, \qquad R = \left(\frac{\theta^2 \|b\|_2^2 - \|a\|_2^2}{1 - \theta^2} + \|x_0\|_2^2 \right)^{1/2}.$$

2.15 *Some sets of probability distributions.* Let x be a real-valued random variable with $\mathbf{prob}(x = a_i) = p_i$, $i = 1, \ldots, n$, where $a_1 < a_2 < \cdots < a_n$. Of course $p \in \mathbf{R}^n$ lies in the standard probability simplex $P = \{p \mid \mathbf{1}^T p = 1, \ p \succeq 0\}$. Which of the following conditions are convex in p? (That is, for which of the following conditions is the set of $p \in P$ that satisfy the condition convex?)

(a) $\alpha \leq \mathbf{E} f(x) \leq \beta$, where $\mathbf{E} f(x)$ is the expected value of $f(x)$, *i.e.*, $\mathbf{E} f(x) = \sum_{i=1}^n p_i f(a_i)$. (The function $f : \mathbf{R} \to \mathbf{R}$ is given.)

(b) $\mathbf{prob}(x > \alpha) \leq \beta$.

(c) $\mathbf{E} |x^3| \leq \alpha \mathbf{E} |x|$.

(d) $\mathbf{E} x^2 \leq \alpha$.

(e) $\mathbf{E} x^2 \geq \alpha$.

(f) $\mathbf{var}(x) \leq \alpha$, where $\mathbf{var}(x) = \mathbf{E}(x - \mathbf{E} x)^2$ is the variance of x.

(g) $\mathbf{var}(x) \geq \alpha$.

(h) $\mathbf{quartile}(x) \geq \alpha$, where $\mathbf{quartile}(x) = \inf\{\beta \mid \mathbf{prob}(x \leq \beta) \geq 0.25\}$.

(i) $\mathbf{quartile}(x) \le \alpha$.

Solution. We first note that the constraints $p_i \ge 0$, $i = 1, \ldots, n$, define halfspaces, and $\sum_{i=1}^n p_i = 1$ defines a hyperplane, so P is a polyhedron.

The first five constraints are, in fact, linear inequalities in the probabilities p_i.

(a) $\mathbf{E}\, f(x) = \sum_{i=1}^n p_i f(a_i)$, so the constraint is equivalent to two linear inequalities

$$\alpha \le \sum_{i=1}^n p_i f(a_i) \le \beta.$$

(b) $\mathbf{prob}(x \ge \alpha) = \sum_{i:\, a_i \ge \alpha} p_i$, so the constraint is equivalent to a linear inequality

$$\sum_{i:\, a_i \ge \alpha} p_i \le \beta.$$

(c) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i(|a_i^3| - \alpha |a_i|) \le 0.$$

(d) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i a_i^2 \le \alpha.$$

(e) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i a_i^2 \ge \alpha.$$

The first five constraints therefore define convex sets.

(f) The constraint

$$\mathbf{var}(x) = \mathbf{E}\, x^2 - (\mathbf{E}\, x)^2 = \sum_{i=1}^n p_i a_i^2 - (\sum_{i=1}^n p_i a_i)^2 \le \alpha$$

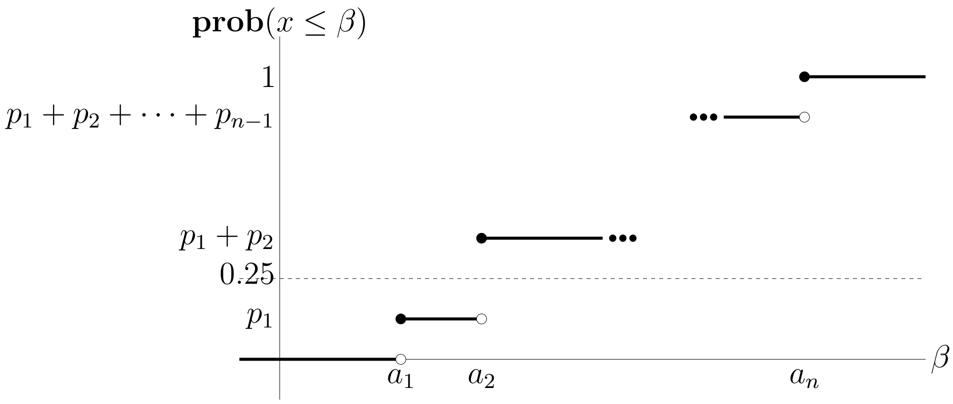
is not convex in general. As a counterexample, we can take $n = 2$, $a_1 = 0$, $a_2 = 1$, and $\alpha = 1/5$. $p = (1, 0)$ and $p = (0, 1)$ are two points that satisfy $\mathbf{var}(x) \le \alpha$, but the convex combination $p = (1/2, 1/2)$ does not.

(g) This constraint is equivalent to

$$-\sum_{i=1}^n a_i^2 p_i + (\sum_{i=1}^n a_i p_i)^2 = -b^T p + p^T A p \le -\alpha$$

where $b_i = a_i^2$ and $A = aa^T$. This defines a convex set, since the matrix aa^T is positive semidefinite.

Let us denote **quartile**$(x) = f(p)$ to emphasize it is a function of p. The figure illustrates the definition. It shows the cumulative distribution for a distribution p with $f(p) = a_2$.



(h) The constraint $f(p) \geq \alpha$ is equivalent to

$$\textbf{prob}(x \leq \beta) < 0.25 \text{ for all } \beta < \alpha.$$

If $\alpha \leq a_1$, this is always true. Otherwise, define $k = \max\{i \mid a_i < \alpha\}$. This is a fixed integer, independent of p. The constraint $f(p) \geq \alpha$ holds if and only if

$$\textbf{prob}(x \leq a_k) = \sum_{i=1}^{k} p_i < 0.25.$$

This is a strict linear inequality in p, which defines an open halfspace.

(i) The constraint $f(p) \leq \alpha$ is equivalent to

$$\textbf{prob}(x \leq \beta) \geq 0.25 \text{ for all } \beta \geq \alpha.$$

Here, let us define $k = \max\{i \mid a_i \leq \alpha\}$. Again, this is a fixed integer, independent of p. The constraint $f(p) \leq \alpha$ holds if and only if

$$\textbf{prob}(x \leq a_k) = \sum_{i=1}^{k} p_i \geq 0.25.$$

If $\alpha \leq a_1$, then no p satisfies $f(p) \leq \alpha$, which means that the set is empty. Thus, the constraint $f(p) \leq \alpha$ is a linear inequality on p.

2.28 *Positive semidefinite cone for* $n = 1,\ 2,\ 3$. Give an explicit description of the positive semidefinite cone $\mathbf{S}_+^n$, in terms of the matrix coefficients and ordinary inequalities, for $n = 1,\ 2,\ 3$. To describe a general element of $\mathbf{S}^n$, for $n = 1,\ 2,\ 3$, use the notation

$$x_1, \qquad \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}, \qquad \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix}.$$

Solution. A symmetric matrix X is positive semidefinite if and only if all principal minors (determinants of symmetric submatrices) are nonnegative. For $n = 1$ the condition is just $x_1 \geq 0$. For $n = 2$ the condition is

$$x_1 \geq 0, \qquad x_3 \geq 0, \qquad x_1x_3 - x_2^2 \geq 0.$$

For $n = 3$ the condition is

$$x_1 \geq 0, \qquad x_4 \geq 0, \qquad x_6 \geq 0, \qquad x_1x_4 - x_2^2 \geq 0, \qquad x_4x_6 - x_5^2 \geq 0, \qquad x_1x_6 - x_3^2 \geq 0$$

and

$$x_1x_4x_6 + 2x_2x_3x_5 - x_1x_5^2 - x_6x_2^2 - x_4x_3^2 \geq 0.$$

2.36 *Euclidean distance matrices.* Let $x_1, \ldots, x_n \in \mathbf{R}^k$. The matrix $D \in \mathbf{S}^n$ defined by $D_{ij} = \|x_i - x_j\|_2^2$ is called a *Euclidean distance matrix*. It satisfies some obvious properties such as $D_{ij} = D_{ji}$, $D_{ii} = 0$, $D_{ij} \geq 0$, and (from the triangle inequality) $D_{ik}^{1/2} \leq D_{ij}^{1/2} + D_{jk}^{1/2}$. We now pose the question: When is a matrix $D \in \mathbf{S}^n$ a Euclidean distance matrix (for some points in $\mathbf{R}^k$, for some k)? A famous result answers this question: $D \in \mathbf{S}^n$ is a Euclidean distance matrix if and only if $D_{ii} = 0$ and $x^TDx \leq 0$ for all x with $\mathbf{1}^Tx = 0$.

Show that the set of Euclidean distance matrices is a convex cone. Find the dual cone.

Solution. The set of Euclidean distance matrices in $\mathbf{S}^n$ is a closed convex cone because it is the intersection of (infinitely many) halfspaces defined by the following homogeneous inequalities:

$$e_i^TDe_i \leq 0, \qquad e_i^TDe_i \geq 0, \qquad x^TDx = \sum_{j,k} x_jx_kD_{jk} \leq 0,$$

for all $i = 1, \ldots, n$, and all x with $\mathbf{1}^Tx = 0$.

It follows that dual cone is given by

$$K^* = \mathbf{conic}(\{-xx^T \mid \mathbf{1}^Tx = 0\} \bigcup \{e_1e_1^T, -e_1e_1^T, \ldots, e_ne_n^T, -e_ne_n^T\}).$$

This can be made more explicit as follows. Define $V \in \mathbf{R}^{n\times(n-1)}$ as

$$V_{ij} = \begin{cases} 1 - 1/n & i = j \\ -1/n & i \neq j. \end{cases}$$

The columns of V form a basis for the set of vectors orthogonal to $\mathbf{1}$, *i.e.*, a vector x satisfies $\mathbf{1}^Tx = 0$ if and only if $x = Vy$ for some y. The dual cone is

$$K^* = \{VWV^T + \mathbf{diag}(u) \mid W \preceq 0, u \in \mathbf{R}^n\}.$$