Intersections of linear subspaces with the closed positive orthant

Wim Couwenberg

July 28, 2008

Abstract

The closed positive orthant in $\mathbb{R}^n$ is the set of all vectors that only have non-negative coordinates. Equip $\mathbb{R}^n$ with the standard inner product. We show that for any linear subspace V in $\mathbb{R}^n$ either V or its orthogonal complement $V^\perp$ contains a non-zero vector in the closed positive orthant.

If v is a vector in $\mathbb{R}^n$ and $j \in \{1, \ldots, n\}$ we write v_j for the j-th coordinate of v. The inner product between two vectors v and w is denoted by (v, w) and the norm of a vector v by $|v|$, so $|v|^2 = (v, v)$. We call v *non-negative* if all coordinates of v are non-negative. The set of all non-negative unit vectors is denoted by P:

$$P = \{v \mid v \in \mathbb{R}^n \mid v \text{ is non-negative and } |v| = 1\}. \tag{1}$$

Theorem 1 *Let V be a linear subspace of $\mathbb{R}^n$ and let $V^\perp$ be its orthogonal complement. Then either V or $V^\perp$ contains a non-zero non-negative vector.*

Let π_V be the orthogonal projection of $\mathbb{R}^n$ onto V. Then any vector v can be decomposed as $v = \pi_V v + (v - \pi_V v)$ with components in V and $V^\perp$. We will use the fact that π_V is self-adjoint and idempotent and in particular

$$(\pi_V v, \pi_V w) = (\pi_V v, w) = (v, \pi_V w) \tag{2}$$

for any pair of vectors v and w. The squared distance $d(v)$ of a vector $v \in \mathbb{R}^n$ to the space V is given by

$$d(v) = |v - \pi_V v|^2 = |v|^2 - |\pi_V v|^2. \tag{3}$$

The squared distance d is homogeneous of degree two and if v and w are orthogonal vectors then

$$d(v + w) = d(v) + d(w) - 2(\pi_V v, w). \tag{4}$$

Let $x \in P$ be a vector where d attains a minimum on the compact set P. Let $J \subseteq \{1, \ldots, n\}$ be the set of coordinate indices j for which $x_j = 0$. Hence

$x_j > 0$ for all coordinate indices $j \notin J$, since x is non-negative. Define the linear subspace W of $\mathbb{R}^n$ by

$$W = \{v \mid v \in \mathbb{R}^n \mid v_j = 0 \text{ for all } j \in J\}. \tag{5}$$

By definition x is contained in W. If $v \in x^\perp \cap W$ and $\lambda \in \mathbb{R}$ is a scalar then we write $v_\lambda = x + \lambda v$, which is again contained in W. Since v is perpendicular to x, the square norm of v_λ reduces to

$$|v_\lambda|^2 = 1 + \lambda^2 |v|^2 \geq 1. \tag{6}$$

Moreover, for each $v \in x^\perp \cap W$ there is an $\epsilon > 0$ such that v_λ is non-negative if $|\lambda| < \epsilon$. Combining this with the fact that $d(x)$ is a minimum of d on P and $|v_\lambda| \geq 1$ we conclude that

$$d(x) \leq d(v_\lambda) = d(x) + \lambda^2 d(v) - 2\lambda(\pi_V x, v) \tag{7}$$

for all scalars λ such that $|\lambda| < \epsilon$. This implies that the linear term in λ in this expression must vanish, so $(\pi_V x, v) = 0$ for all $v \in x^\perp \cap W$. Combining $\pi_V x$ with a well chosen multiple of x we find that

$$(\pi_V x - |\pi_V x|^2 x, x) = |\pi_V x|^2 - |\pi_V x|^2 = 0 \tag{8}$$

and hence $\pi_V x - |\pi_V x|^2 x \in W^\perp$ since W is spanned by x and $x^\perp \cap W$. In other words, for each coordinate index $j \notin J$ we have

$$(\pi_V x)_j = |\pi_V x|^2 x_j. \tag{9}$$

At this point we make the observation that if $J = \emptyset$ then $\pi_V x$ is a scalar multiple of x and hence equal to x or 0 since π_V is idempotent. If $\pi_V x = 0$ then $x = x - 0 = x - \pi_V x \in V^\perp$. So either V or $V^\perp$ contains $x \in P$ which is a non-zero non-negative vector.

Now assume that $\pi_V x$ is not a multiple of x and $J \neq \emptyset$. Take some $j \in J$. Denoting the j-th standard basis vector of $\mathbb{R}^n$ by e_j and taking any scalar $\lambda \geq 0$ it is immediate that $x + \lambda \mathrm{e}_j$ is non-negative and has square norm $1 + \lambda^2 \geq 1$. Repeating the argument that we followed for v_λ we now find that

$$d(x) \leq d(x + \lambda \mathrm{e}_j) = d(x) + \lambda^2 d(\mathrm{e}_j) - 2\lambda(\pi_V x)_j \tag{10}$$

for all scalars $\lambda \geq 0$ and hence

$$(\pi_V x)_j \leq 0 \tag{11}$$

for all $j \in J$ with a strict inequality for at least one index in J since $\pi_V x$ is not a multiple of x. So $|\pi_V x|^2 > 0$ and therefore $(\pi_V x)_j > 0$ for every coordinate index $j \notin J$. This implies that V has a positive distance to P since its closest vector to P (namely $\pi_V x$) has both negative and positive coordinates. However, if we define $y = x - \pi_V x$ (so y is the component of x in $V^\perp$) then y is non-zero and non-negative. This can be easily checked by what we deduced so far: $y_j = -(\pi_V x)_j \geq 0$ for $j \in J$ and $y_j = (1 - |\pi_V x|^2)x_j = d(x)x_j \geq 0$ for $j \notin J$. This concludes the proof.

Note that the theorem does not generalize to more than two orthogonal components. As an example, consider the three vectors

$$(-1, 2, 2),\ (2, -1, 2),\ (2, 2, -1) \in \mathbb{R}^3. \tag{12}$$

These vectors are mutually orthogonal but the three lines spanned by each of these vectors only intersect the closed positive octant of $\mathbb{R}^3$ in the origin.