Tutorial Sheet

All matrices are assumed to have real entries. Unless otherwise mentioned

- A denotes a matrix of size $m \times n$;
- x a $n \times 1$ matrix of indeterminantes; and
- b a $m \times 1$ matrix (in other words, a vector in $\mathbb{R}^m$)

(1) Eliminate x_1 by the Fourier-Motzkin proccess from the following system of five inequalities: $x_1 + x_2 \geqslant 2$; $x_2 \geqslant 1$; $2x_2 - x_1 \leqslant 4$; $x_1 - x_2/3 \leqslant 2$; $x_1 + x_2 \leqslant 5$.

(2) The purpose of this exercise is to exploit Fourier-Motzkin elimination to prove the following version of Farkas' Lemma[1]:

> $Ax \leqslant b$ has a solution if and only if: for every $y \in \mathbb{R}^m$ such that $y \geqslant 0$ and $y^t A = 0$, we have $0 \leqslant y^t b$.

The "only if" part of the statement is straight forward: check this.

For the "if" part of the statement, we prove the contrapositive: assuming that $Ax \leqslant b$ has no solution, we construct a $y \in \mathbb{R}^m$ such that $y \geqslant 0$, $y^t A = 0$, and $y^t b = -1$. And for this, we proceed by induction on the number n of variables.

(a) Base case of the induction proof: Assume $n = 1$. In this case, construct the desired y directly with bare hands.

(b) Induction step: Use Fourier-Motzkin elimination on $Ax \leqslant b$ to get a system $\tilde{A}\tilde{x} \leqslant \tilde{b}$ of linear inequalities in $n-1$ variables ($\tilde{x} = (x_2, \ldots, x_n)^t$) that has no solution. By the induction hypothesis, there exists $\tilde{y} \in \mathbb{R}^{\tilde{m}}$ such that $\tilde{y} \geqslant 0$, $\tilde{y}^t A = 0$ and $\tilde{y}^t b = -1$. Use this $\tilde{y}$ and the relation between the two systems $Ax \leqslant b$ and $\tilde{A}\tilde{x} \leqslant \tilde{b}$ to get the desired $y \in \mathbb{R}^m$. $\square$

(3) The purpose of this exercise is to derive the following version of Farkas' Lemma from the version in the previous exercise:

> The system $Ax = b$, $x \geqslant 0$ has a solution if and only if: for every $y \in \mathbb{R}^m$ such that $0 \leqslant y^t A$, we have $0 \leqslant y^t b$.

Once again, the "only if" part of the statement is straight forward: check this.

For the "if" part of the statement, we concentrate, once again, on the contrapositive: assuming that $Ax = b$, $x \geqslant 0$ has no solution, we construct a $y \in \mathbb{R}^m$ such that $0 \leqslant y^t A$ and $-1 = y^t b$. And for this, we apply the "only if" (or, more correctly, its contrapositive) of the previous version to the system $Ax \leqslant b$, $-Ax \leqslant -b$, $-x \leqslant 0$ to obtain a $\tilde{y} \in \mathbb{R}^{2m+n}$ with $\tilde{y} \geqslant 0$, $\tilde{y}^t A = 0$, and $-1 = \tilde{y}^t b$. Extract from this $\tilde{y}$ an appropriate $y \in \mathbb{R}^m$ such that $y \geqslant 0$, $0 \leqslant y^t A$, but $y^t b = -1$. $\square$

[1] There are other proofs of Farkas' Lemma.

(4) We prove yet another version of Farkas' Lemma in the following exercise:

> The system $Ax = 0$, $x \geqslant 0$ has a non-trivial solution if and only if: there does not exist $y \in \mathbb{R}^m$ such that $y^t A > 0$ (equivalently $y^t A < 0$).

Note that the system always admits the trivial solution. The "only if" part of the statement is, once again, straight forward: check this.

For the "if" part of the statement, we prove, once again, the contrapositive. So suppose that the system $Ax = 0$, $x \geqslant 0$ has no non-trivial solution. We modify it in n different ways to get n different systems, none of which has a solution. Namely, for every j, $1 \leqslant j \leqslant n$, consider $Ax = 0$, $e_j^t x = 1$, $x \geqslant 0$. Now, by the contrapositive of the "only if" portion of the previous version (the one in the previous item), we get, for every j, a $y_j \in \mathbb{R}^m$ such that $y_j^t A \geqslant e_j^t$. Put $y = \sum_{j=1}^n y_j$. Then $y^t A > 0$. □

(5) Let P be a $n \times n$ stochastic matrix, which just means that all the entries of P are non-negative and that all the row sums of P are equal to 1. Prove that there exists a non-trivial vector $\pi \in \mathbb{R}^n$ such that $\pi \geqslant 0$ and $\pi^t P = \pi^t$. (Hint: Rewrite $\pi^t P = \pi^t$ and $(I - P^t)\pi = 0$. Putting $A = (I - P^t)$, we see that what we want is a non-trivial solution to $Ax = 0$, $x \geqslant 0$. By the version of Farkas' Lemma in the previous item, it is enough to show that $y^t A$ is never strictly positive (for any $y \in \mathbb{R}^n$). By way of contradiction, suppose there is a $y \in \mathbb{R}^n$ such that $y^t A > 0$. This means that $y_i > \sum_{j=1}^n p_{ij} y_j$ for each i, $1 \leqslant i \leqslant n$. Now take i to be such that y_i has the least possible value (among all the entries of y). Then, since $\sum_{j=1}^n p_{ij} = 1$ with p_{ij} all being non-negative, this leads to a contradiction.) □

We consider solution sets first of homogeneous systems $Ax \geqslant 0$ and then of inhomogeneous systems $Ax \geqslant b$.

Case of a Homogeneous System $Ax \geqslant 0$**.** The solution set is a convex cone in $\mathbb{R}^n$. Convex cones that arise in such fashion (for some choice of the matrix A) are called polyhedral (for the reason that they are finite intersections of half-spaces). Thus let us denote the solution set $\{x \in \mathbb{R}^n \mid Ax \geqslant 0\}$ by $\mathrm{PHC}(A)$: "PHC" stands for "polyhedral cone".

Our goal is to show that $\mathrm{PHC}(A)$ is finitely generated or polytopal in the following sense. Given an $n \times p$ matrix D, we denote by $\mathrm{FGC}(D)$ the convex cone $\{Dy \mid 0 \leqslant y \in \mathbb{R}^p\}$: "FGC" stands for "finitely generated cone". Note that $\mathrm{FGC}(D)$ consists of all non-negative linear combinations of the (finitely many) columns of D.

The following lemma gives examples of such finitely generated cones:

[l:wposfg]

Lemma 1. *The intersection of a vector subspace of* $\mathbb{R}^n$ *with the positive orthant is a finitely generated cone. More precisely, given a vector subspace* W *of* $\mathbb{R}^n$*, we can find a* $n \times p$ *matrix* D *such that* $W \cap \{x \in \mathbb{R}^n \mid x \geqslant 0\} = \mathrm{FGC}(D)$.

A proof of the lemma is discussed in Tutorial Sheet 2.

The duals and double duals of polytopal and polyhedral cones. Given a cone K in $\mathbb{R}^n$, its dual $K^\star$ is defined by $K^\star := \{y \in \mathbb{R}^n \mid y^t x \geqslant 0 \text{ for all } x \text{ in } K\}$.

[p:dual]

Proposition 2. *(1)* $\mathrm{FGC}(D)^\star = \mathrm{PHC}(D^t)$.
(2) $\mathrm{PHC}(A)^\star = \mathrm{FGC}(A^t)$.
(3) $K^{\star\star} = K$ *if* K *is polytopal or polyhedral.*

Proof: Item (1) is straight forward, while (2) is a restatement of Farkas' Lemma (version 3). As for (3), using (1) and (2), we have: $\mathrm{FGC}(D)^{\star\star} = \mathrm{PHC}(D^t)^\star = \mathrm{FGC}(D)$ and $\mathrm{PHC}(A)^{\star\star} = \mathrm{FGC}(A^t)^\star = \mathrm{PHC}(A)$. □

Polyhedral cones are polytopal and vice versa (Minkowski-Weyl Theorem).

[t:phispt]

Theorem 3. *Given an* $m \times n$ *matrix* A*, there exists an* $n \times p$ *matrix* D *(for some choice* p*) such that* $\mathrm{PHC}(A) = \mathrm{FGC}(D)$*. In other words, every polyhedral cone is polytopal.*

Proof: First consider the null space $\{x \mid Ax = 0\}$. Choose a basis for this space. Let D_0 be the matrix whose columns are these basis elements and also their negatives. Then any element of the null space is a non-negative linear combination of the columns of D_0.

Now consider the image W in $\mathbb{R}^m$ of the linear map $x \mapsto Ax$. By Lemma 1, there is a $m \times q$ matrix F such that $FGC(F) = W \cap \mathbb{R}^m_{\geqslant 0}$. Since the columns of F belong to the image of A, we can find an $n \times q$ matrix D_1 such that $AD_1 = F$.

Put $D := (D_0|D_1)$. We claim that $\mathrm{PHC}(A) = \mathrm{FGC}(D)$. Since the columns of D belong to $\mathrm{PHC}(A)$ by construction, it is clear that $\mathrm{PHC}(A) \supseteq \mathrm{FGC}(D)$. For the other inclusion, suppose x belongs to $\mathrm{PHC}(A)$. Then $Ax \geqslant 0$ and so there exists $y \in \mathbb{R}^q$, $y \geqslant 0$ such that $Ax = Fy$. Substituting $F = AD_1$ into this, we see that $A(x - D_1 y) = 0$, and so $x - D_1 y$ belongs to $\mathrm{FGC}(D_0)$. But then $x = (x - D_1 y) + D_1 y$ belongs to $\mathrm{FGC}(D_0|D_1) = \mathrm{FGC}(D)$. □

Corollary 4. *Given an* $n \times p$ *matrix* D*, there exists an* $m \times n$ *matrix* A *(for some choice* m*) such that* $\mathrm{FGC}(D) = \mathrm{PHC}(A)$*. In other words, every polytopal cone is polyhedral.*

Proof: By Proposition 2, we have $\mathrm{FGC}(D) = \mathrm{FGC}(D)^{\star\star} = \mathrm{PHC}(D^t)^\star$. By Theorem 3, there exists an $n \times m$ matrix A^t (for some m) such that $\mathrm{PHC}(D^t) = \mathrm{FGC}(A^t)$. Putting these together and using Proposition 2 again, we get $\mathrm{FGC}(D) = \mathrm{PHC}(D^t)^\star = \mathrm{FGC}(A^t)^\star = \mathrm{PHC}(A)$. $\square$

Corollary 5. *The intersection of two polytopal cones is polytopal. The sum of two polyhedral cones is polyhedral.*

Proof: The intersection of two polyhedral cones is evidently polyhedral. The sum of two polytopal cones is evidently polytopal. Since polytopal cones are polyhedral and vice versa, the corollary follows. $\square$

The case of a general inhomogeneous system $Ax \geqslant b$. The solution set $\{x \mid Ax \leqslant b\}$ is a convex polyhedral object in $\mathbb{R}^n$ (polyhedral meaning that it is a finite intersection of half-spaces). To get at the structure of such a set in general, the strategy we follow is to homogenize the system by the addition of one more variable, say t, and realize the solution set as the intersection of the solution set of the homogeneous system with the hyperplane $t = 1$.

So consider the homogeneous system $Ax - tb \geqslant 0$, $t \geqslant 0$, and let $\tilde{S}$ be its solution set in $\mathbb{R}^{n+1}$. The solution set S of the original system $Ax \geqslant b$ is related to $\tilde{S}$ thus: $S = \{x \in \mathbb{R}^n \mid (x, 1) \in \tilde{S}\}$. Now, as already seen above (Theorem 3), there exists a $(n + 1) \times p$ matrix, say $\tilde{D}$, such that $\tilde{S} = \mathrm{FGC}(\tilde{D}) = \{\tilde{D}y \mid 0 \leqslant y \in \mathbb{R}^p\}$.

Since the inequality $t \geqslant 0$ is part of our system, it follows that the last coordinate (that is, the t-coordinate) of each of the columns of $\tilde{D}$ is non-negative. Put $\tilde{D} = (\tilde{D}_1 \mid \tilde{D}_0)$, where $\tilde{D}_1$ is the collection of all those columns of $\tilde{D}$ where the last coordinate is strictly positive, and $\tilde{D}_0$ is the collection of those columns where that coordinate is zero. We may further assume that the value of the last coordinate in each column of $\tilde{D}_1$ equals 1, by appropriate scaling.

Let D_1 (respectively D_0) be the sub-matrix of $\tilde{D}_1$ (respectively $\tilde{D}_0$) consisting of the first n rows. We claim that S admits the following description:

[t:strinhom]

Theorem 6. *The solution set S of the system $Ax \geqslant b$ is the sum of two convex objects, namely the convex hull of the columns of D_1 and the convex cone consisting of the non-negative linear combinations of the columns of D_0.*

To justify this claim (in other words, to finish the proof of the theorem), it is convenient to introduce one more piece of notation. Let us call a vector y in $\mathbb{R}^q$ (where q is some positive integer) stochastic if each of the components of y is non-negative and the sum of the components is 1. Given a $n \times q$ matrix E, let us denote the convex hull $\{Ey \mid y \in \mathbb{R}^q \text{ is stochastic}\}$ of the columns of E by C-Hull(E). With this in place, the theorem can be stated thus:

$$S := \{x \in \mathbb{R}^n \mid Ax \geqslant b\} = \text{C-Hull}(D_1) + \mathrm{FGC}(D_0)$$

We now prove this equality. First suppose $x \in S$. Then $(x, 1) \in \tilde{S}$, and so $(x, 1) = (\tilde{D}_1 \mid \tilde{D}_0)y$ for some $y \geqslant 0$. Letting p_1 denote the number of columns of $\tilde{D}_1$, and equating the last coordinate on both sides we get $1 = \sum_{i=1}^{p_1} y_i$. Equating the other co-ordinates, we have $x = \sum_{i=1}^{p_1}(D_1)_i y_i + \sum_{i=p_1+1}^{p}(D_0)_i y_i$. Of the two summations on the right hand side of this equation, the first belongs to C-Hull(D_1) and the second to $\mathrm{FGC}(D_0)$.

Now suppose x belongs to C-Hull(D_1) + $\mathrm{FGC}(D_0)$, or in other words $x = \sum_{i=1}^{p_1}(D_1)_i y_i + \sum_{i=p_1+1}^{p}(D_0)_i y_i$ for some $y \in \mathbb{R}^p$ with $y_i \geqslant 0$ for all i and $\sum_{i=1}^{p_1} y_i = 1$. Then $(x, 1) = (\tilde{D}_1 \mid \tilde{D}_0)y$ and so $(x, 1) \in \tilde{S}$ and $x \in S$. $\square$

Remark 7. The cone FGC(D_0) is precisely the solution set of the homogeneous system $Ax \geqslant 0$. Indeed, for a vector $x \in \mathbb{R}^n$,we have $Ax \geqslant 0 \Leftrightarrow (x, 0) \in \tilde{S} \Leftrightarrow (x, 0)$ is a non-negative linear combination of the columns of $\tilde{D}_0$ (since the last co-ordinates of the columns of $\tilde{D}_1$ are all 1, they cannot appear in the expression for $(x, 0)$ as a non-negative linear combination of the columns of $\tilde{D}$). Finally, this last condition on $(x, 0)$ is equivalent to the condition that x belong to FGC(D_0).

Tutorial Sheet 2

The notation and conventions of Tutorial Sheet 1 continue to hold.

(1) Prove or disprove: The system $Ax \leqslant b$, $x \geqslant 0$ has a solution if and only if: for every $y \geqslant 0$ in $\mathbb{R}^m$, if $y^t A \geqslant 0$ then $y^t b \geqslant 0$.

(2) This is an exercise in reading comprehension! Read critically the proof below of Lemma 1.
Proof: Proceed by induction on the dimension of W. First suppose that $\dim W = 1$. Let w be a non-zero vector in W. If w (or $-w$) belongs to the positive orthant, we choose D to be the $n \times 1$ matrix whose sole column is w (or $-w$). If not, we could take D to be the zero $n \times 1$ matrix.

Now suppose $\dim W \geqslant 2$. Let us denote the positive orthant of $\mathbb{R}^n$ by $\mathbb{R}^n_{\geqslant 0}$, and, for any subspace V in $\mathbb{R}^n$, its intersection with $\mathbb{R}^n_{\geqslant 0}$ by $V_{\geqslant 0}$. Let I be the subset of those i, $1 \leqslant i \leqslant n$, for which $W_{\geqslant 0}$ has a vector whose i^{th} component is non-zero. For i in I, let W_i be the subspace of W consisting of vectors in W whose i^{th} component is zero. Then $\dim W(i) < \dim W$ (why?). By the induction hypothesis $W(i)_{\geqslant 0}$ is finitely generated, say by the columns of a $n \times p_i$ matrix D_i.

Let w be a vector in $W_{\geqslant 0}$ such that, for each $i \in I$, the i^{th} component of w is non-zero. (Why does such a w exist? For each $i \in I$, choose $w_i \in W_{\geqslant 0}$ such that the i^{th} component of w_i is positive. The vector $w = \sum_{i \in I} w_i$ has the desired property.) Let D be the matrix (with n rows) whose first column is w and whose later columns are constructed by putting together the matrices D_i, $i \in I$.

We claim that $W_{\geqslant 0} =$ FGC(D). Since the columns of D all belong to $W_{\geqslant 0}$ by construction, it is clear that $W_{\geqslant 0} \supseteq$ FGC(D). To prove the reverse inclusion, let x be in $W_{\geqslant 0}$. If, for any $i \in I$, the i^{th} component of x is zero, then x belongs to $W(i)_{\geqslant 0}$ and so is generated by the columns of D_i (and hence also of D). So, suppose that is not the case. Consider $x - tw$, where t is a non-negative real parameter. For small values of t, $x - tw$ belongs to $W_{\geqslant 0}$. In fact, there is a largest value of t for which $x - tw$ still belongs to $W_{\geqslant 0}$. Fix t to be that value. Then $x - tw$ belongs to $W(i)_{\geqslant 0}$ for some appropriately chosen $i \in I$, and so is a non-negative linear combination of the columns of D_i. Thus $x = tw + (x - tw)$ is a non-negative linear combination of the columns of D. □