Relaxations

In this section we approximate difficult optimization problems by a semi-definite program. Usually the solution of the SDP furnishes a lower bound on the optimal value of the original problem. In some cases we can give performance guarantees and recover a suboptimal solution of the original problem from the optimal solution of the SDP.

10.1 Max-Cut

Consider again the Max-Cut problem. We are given a graph $G$ with edge weights $w_{ij}$, $i,j = 1, \ldots, n$. The problem consists in partitioning the vertices $1, \ldots, n$ of $G$ into two disjoint subsets $S, T$ such that the weight $\sum_{i \in S, j \in T} w_{ij}$ of the resulting cut is maximized. In contrast to the Min-Cut problem with nonnegative weights this problem is NP-complete.

We represent the partition into subsets $S, T$ by a vector $x \in \{-1, +1\}^n$, for any partition the weight of the cut is then given by the matrix scalar product $\langle A, W \rangle$, where $W = (W_{ij})$ is the matrix of the edge weights and $A = \frac{1}{4}(1 - xx^T) = \frac{1}{4}(1 - X)$. The problem can then be written as

$$\max_{X \in S^n} \frac{1}{4} \langle W, 1 - X \rangle : \ \text{diag}(X) = 1, \ \text{rk} X = 1.$$  

(1)

This would be a semi-definite program, if there were not the rank constraint. By dropping the rank constraint we obtain the semi-definite relaxation

$$\max_{X \in S^n} \frac{1}{4} \langle W, 1 - X \rangle : \ \text{diag}(X) = 1.$$  

(2)

Clearly the optimal value $c_{MC}^{opt}$ of (1) is upper bounded by the optimal value $c_{SR}^{opt}$ of (2), because the feasible set of the latter is overbounding the former.

We shall now employ a randomized rounding procedure to generate suboptimal solutions of the original problem from the solution of the relaxed problem, due to Goemans and Williamson [2]. Let $X_{SR}^{opt}$ be the maximizer in (2), and let $k$ be its rank. Then $X$ is the Gramian of $n$ unit norm vectors $f_1, \ldots, f_n \in \mathbb{R}^k$, which can be computed as columns of the $k \times n$ matrix $F$ in the factorization $X_{SR}^{opt} = F^TF$. Let now $\xi \in \mathbb{R}^k$ be a uniformly distributed random unit norm vector. For each such $\xi$, we generate a cut by setting

$$S_\xi = \{i \mid \langle \xi, f_i \rangle \geq 0\}, \quad T_\xi = \{i \mid \langle \xi, f_i \rangle < 0\}.$$  

(3)

In other words, the hyperplane which is normal to $\xi$ separates the vectors $f_i$ into two subsets, which will serve as the partition defining the cut.

Let us compute the expectation of the weight $c_{\xi}^{subopt}$ of this cut. The probability that $\xi$ separates two vectors $f_i, f_j$ depends only on the angle $\theta_{ij} = \arccos\langle f_i, f_j \rangle = \arccos(X_{SR}^{opt})_{ij}$ between $f_i$ and $f_j$ and increases linearly from 0 for $\theta_{ij} = 0$ to 1 for $\theta_{ij} = \pi$. The expectation then equals

$$\mathbb{E}_{c_{\xi}^{subopt}} = \frac{1}{2\pi} \sum_{i,j=1}^n w_{ij} \theta_{ij} = \frac{1}{2\pi} \sum_{i,j=1}^n w_{ij} \arccos(X_{SR}^{opt})_{ij}.$$  

Let us now assume that the weights $w_{ij}$ are nonnegative. We have the inequality

$$\alpha \cdot \frac{1}{4}(1 - (X_{SR}^{opt})_{ij}) \leq \frac{1}{2\pi} \arccos(X_{SR}^{opt})_{ij},$$

where $\alpha = \frac{2}{\pi} \min_{y \in (-1,1)} \frac{\arccos y}{1 - y} = \frac{2}{\pi} \min_{\theta \in [0,\pi]} \frac{\theta}{1 - \cos \theta} \approx 0.87856$. Multiplying by $w_{ij}$ and summing over $i, j$, we obtain

$$\alpha \cdot c_{SR}^{opt} \leq \mathbb{E}_{c_{\xi}^{subopt}} \leq c_{MC}^{opt} \leq c_{SR}^{opt}.$$  

Thus the upper bound $c_{SR}^{opt}$ is not larger than $\alpha^{-1} \approx 1.138217$ times the optimal value $c_{MC}^{opt}$ of the original problem. On the other hand, by repeatedly generating random cuts we obtain sub-optimal objective values which are not smaller than $\mathbb{E}_{c_{\xi}^{subopt}} - \varepsilon$ for $\varepsilon$ as small as desired. The expectation is in turn not smaller than $\alpha$ times the optimal objective value.

Relaxation (2) hence has a guaranteed performance.
10.2 Nesterov’s $\frac{\pi}{2}$ theorem

Let us have another look at the Max-Cut problem. The difference between this problem and its relaxation is that in (1) the matrix variable $X$ takes values at the vertices of the Max-Cut polytope $\mathcal{MC}$, defined as the convex hull of the set $\{xx^T \mid x \in \{-1, +1\}^n\}$, and in (2) it takes values in the set

$$\mathcal{SR} = \{ X \in S^n_+ \mid \text{diag}(X) = 1 \}.$$ 

In (1) it is actually not relevant whether we maximize over the extreme points of $\mathcal{MC}$ or over the set $\mathcal{MC}$ itself, because the objective function is linear and will assume its maximum at an extreme point anyway.

We shall discard the factor $\frac{1}{4}$ and the constant term $\langle 1, W \rangle$ in the cost function, and consider the two problems

$$\max_{X \in \mathcal{MC}} \langle W, X \rangle \quad (4)$$

and

$$\max_{X \in \mathcal{SR}} \langle W, X \rangle. \quad (5)$$

Problem (5) is a semi-definite relaxation of problem (4), and the optimal value $c_{\mathcal{SR}}^{\text{opt}}$ of the former is an upper bound for the optimal value $c_{\mathcal{MC}}^{\text{opt}}$ of the latter, because we have the inclusion $\mathcal{MC} \subset \mathcal{SR}$.

For every $X \in \mathcal{SR}$ we now consider the matrix $\tilde{X} = \frac{2}{\pi} \arcsin X$, where the arcsin function is applied element-wise.

**Lemma 10.1.** For every $X \in \mathcal{SR}$ the matrix $\tilde{X}$ is an element of $\mathcal{MC}$.

**Proof.** As in the previous section, let $k$ be the rank of $X$, and let $F \in \mathbb{R}^{k \times n}$ such that $X = F^T F$. Let $f_1, \ldots, f_n \in \mathbb{R}^k$ be the columns of $F$ and let $\xi \in \mathbb{R}^k$ be a uniformly distributed random unit length vector. Let $x_{\xi} \in \{-1, +1\}$ be the vector defined by partition (3), and set $X_{\xi} = x_{\xi} x_{\xi}^T$. Then $X_{\xi}$ is a random vertex of the Max-Cut polytope $\mathcal{MC}$. Moreover, we have $(X_{\xi})_{ij} = -1$ with probability $\frac{\arccos X_{\xi,ij}}{\pi}$, and $(X_{\xi})_{ij} = 1$ with probability $1 - \frac{\arccos X_{\xi,ij}}{\pi}$. Therefore

$$E(X_{\xi})_{ij} = 1 - \frac{2}{\pi} \arccos X_{\xi,ij} = \frac{2}{\pi} \arcsin X_{\xi,ij}.$$ 

Hence $E X_{\xi} = \tilde{X}$, and $\tilde{X}$ is a convex combination of the vertices of $\mathcal{MC}$. The proof is concluded by the convexity of $\mathcal{MC}$.

The set

$$\mathcal{T} \mathcal{A} = \{ \frac{2}{\pi} \arcsin X \mid X \in \mathcal{SR} \}$$

Figure 1: Element-wise functions involved in the Goemans-Williamson procedure.
is hence an inner approximation of the Max-Cut polytope, the so-called trigonometric approximation [3]. For \( n \leq 3 \) this relaxation is exact, i.e., \( \mathcal{T}A = \mathcal{MC} \).

The function arcsin has the Taylor series
\[
\arcsin z = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)!!} \frac{z^{2k+1}}{(2k+1)!},
\]
where the double factorial means a product of only odd or only even integers. It follows that the difference \( 2\pi \arcsin z - 2\pi z \) has only nonnegative Taylor coefficients and converges for all \( |z| < 1 \). As a consequence, for every matrix \( X \in \mathcal{SR} \) we have \( \tilde{X} = \frac{2}{\pi} \arcsin X \succeq \frac{2}{\pi} X \).

Suppose now that the weighting matrix \( W \) is itself positive semi-definite. Setting \( X \) equal to the optimal solution \( X_{\text{opt}}^{\mathcal{SR}} \) of the relaxation (5) and taking the matrix scalar product of the preceding inequality with \( W \), we obtain that
\[
\mathbb{E}(W, X_\xi) = \langle W, \tilde{X} \rangle \geq \frac{2}{\pi} \langle W, X_{\text{opt}}^{\mathcal{SR}} \rangle = \frac{2}{\pi} c_{\text{opt}}^{\mathcal{SR}}.
\]

We obtain the following performance guarantee of the semi-definite relaxation [5].

**Theorem 10.2.** Consider problems (4), (5) with a positive semi-definite weighting matrix \( W \) and let \( c_{\text{opt}}^{\mathcal{MC}} \), \( c_{\text{opt}}^{\mathcal{SR}} \) be their optimal values, respectively. Then
\[
\frac{2}{\pi} c_{\text{opt}}^{\mathcal{SR}} \leq c_{\text{opt}}^{\mathcal{MC}} \leq c_{\text{opt}}^{\mathcal{SR}}.
\]

Suboptimal solutions which have at least \( (\frac{2}{\pi} - \varepsilon) \) times the optimal value can be obtained by the described randomized rounding procedure.

### 10.3 Quadratically constrained quadratic programs

We shall now consider a very general class of NP-hard optimization problems, the quadratically constrained quadratic programs (QCQP):

\[
\min_{x \in \mathbb{R}^n} (x^T A_0 x + 2b_0^T x + c_0) : \quad \begin{cases}
  x^T A_i x + 2b_i^T x + c_i = 0, \\
  x^T A'_j x + 2b'_{ij}^T x + c'_j \leq 0, \\
  Cx = d.
\end{cases}
\]  
(6)

Thus the objective function of the problem is quadratic, and the constraints are linear and quadratic.

Let us introduce the vector \( y = (1, x^T)^T \in \mathbb{R}^{n+1} \), whose entries we index from 0 to \( n \), the rank 1 matrix \( Y = yy^T \), and the matrices
\[
A_i = \begin{pmatrix} c_i & b_i^T \\ b_i & A_i \end{pmatrix}, \quad A'_j = \begin{pmatrix} c'_{ij} & b'_{ij}^T \\ b'_{ij} & A'_j \end{pmatrix}, \quad C = (-d, C).
\]

Then our problem can be rewritten as
\[
\min_{Y \in S^{n+1}_+} \langle A_0, Y \rangle : \quad \langle A_i, Y \rangle = 0, \quad \langle A'_j, Y \rangle \leq 0, \quad CY = 0, \quad Y_{00} = 1, \quad \text{rk} \ Y = 1.
\]  
(7)

By dropping the rank constraint we obtain a semi-definite relaxation of the QCQP. Its optimal value lower bounds the optimal value of the original QCQP.

In general nothing can be said about the performance of the semi-definite relaxation.
10.4 Dines’ theorem

We consider the following special cases of QCQPs:
\[
\begin{align*}
\min_{x \in \mathbb{R}^n} x^T A x & : x^T B x = b, \\
\min_{x \in \mathbb{R}^n} x^T A x & : x^T B x \leq b.
\end{align*}
\] (8)

Setting \( X = x x^T \), we can rewrite these problems as
\[
\begin{align*}
\min_{X \in S_n^+} \langle A, X \rangle & : \langle B, X \rangle = b, \ \text{rk} \ X = 1, \\
\min_{X \in S_n^+} \langle A, X \rangle & : \langle B, X \rangle \leq b, \ \text{rk} \ X = 1.
\end{align*}
\] (10) (11)

We have the following result.

**Lemma 10.3.** The semi-definite relaxations of problems (8), (9) which are obtained by dropping the rank constraint from (10), (11), respectively, are exact.

The proof of the lemma relies on the following theorem by Dines [1].

**Theorem 10.4.** Let \( A, B \in S^n \) be real symmetric matrices. Then the set \( \{(x^T A x, x^T B x) \in \mathbb{R}^2 \mid x \in \mathbb{R}^n\} \) is convex.

The set \( \{(x^T A x, x^T B x) \in \mathbb{R}^2 \mid x \in \mathbb{R}^n\} \) is called the numerical range of the pair \((A, B)\). The theorem then says that the numerical range is a convex cone.

**Proof.** (of Lemma 10.3) If the semi-definite relaxations are infeasible, then the original problems are also infeasible. We shall hence assume that the semi-definite relaxations are feasible.

Let \( \bar{X} \in S_n^+ \) be a feasible point, and set \( \bar{b} = \langle B, \bar{X} \rangle \), \( \bar{a} = \langle A, \bar{X} \rangle \). Let \( k = \text{rk} \ \bar{X} \) and factor \( \bar{X} = F F^T \), where the columns of the matrix \( F \in \mathbb{R}^{n \times k} \) are denoted by \( f_1, \ldots, f_k \). Then \( \bar{X} = \sum_{i=1}^k f_i f_i^T \). The points \((f_i^T A f_i, f_i^T B f_i)\) are in the numerical range of \((A, B)\). Since the numerical range is a convex cone, the conic combination
\[
\sum_{i=1}^k (f_i^T A f_i, f_i^T B f_i) = (\langle A, \bar{X} \rangle, \langle B, \bar{X} \rangle) = (\bar{a}, \bar{b})
\]
is also in the numerical range. Hence there exists a vector \( \hat{x} \in \mathbb{R}^n \) such that \((\hat{x}^T A \hat{x}, \hat{x}^T B \hat{x}) = (\bar{a}, \bar{b})\). Therefore problems (10), (11) are equivalent to their counterparts without the rank constraint. \(\square\)

11 Sums of squares and moment relaxations

In this section we deal with optimization problems whose feasible sets are sets of polynomials satisfying certain positivity constraints. The decision variables are hence the coefficient vectors of these polynomials. The vector spaces underlying the optimization problems are thus essentially finite-dimensional function spaces.

References on semi-definite relaxations in polynomial optimization are [4, 6].

11.1 Positive polynomials and sums of squares

**Definition 11.1.** The set \( P_{d,n} \) of positive polynomials in \( n \) variables of degree \( d \) is defined as the set of homogeneous polynomials \( p : \mathbb{R}^n \rightarrow \mathbb{R} \) of degree \( d \) such that \( p(x) \geq 0 \) for all \( x \in \mathbb{R}^n \).

Clearly \( d \) has to be even in order for \( P_{d,n} \) to contain a non-zero element. The set \( P_{d,n} \) is a closed convex cone.

**Definition 11.2.** The set \( \Sigma_{d,n} \) of sum of squares (SOS) polynomials in \( n \) variables of degree \( d \) is defined as the set of homogeneous polynomials \( p : \mathbb{R}^n \rightarrow \mathbb{R} \) of degree \( d \) which can be represented as a finite sum \( p(x) = \sum_k q_k^2(x) \), where \( q_k(x) \) are homogeneous polynomials of degree \( d/2 \) in \( n \) variables.

Clearly \( \Sigma_{d,n} \subset P_{d,n} \).
Theorem 11.3. Let $n, d$ be positive integers, $d$ even. The equality $\Sigma_{d,n} = P_{d,n}$ holds if and only if $\min(d, n) \leq 2$ or if $(d, n) = (4, 3)$.

We shall prove only the cases $d = 2$ and $n = 2$. The case $(d, n) = (4, 3)$ is more involved and has been proven by Hilbert in 1888.

$d = 2$. A homogeneous polynomial of degree 2 is a quadratic form, which can be written as $p(x) = x^T A x$ with $A$ a real symmetric matrix. The polynomial $p$ is in $P_{d,n}$ if and only if the corresponding matrix $A$ is positive semi-definite. This in turn is the case if and only if there exists a matrix $B \in \mathbb{R}^{k \times n}$ such that $A = B^T B$.

On the other hand, $p \in \Sigma_{2,n}$ if and only if $p(x) = \sum_{j=1}^k q_j^2(x)$, where each $q_j$ is a linear homogeneous function, i.e., $q_j(x) = c_j^T x$ for some vector $c_j \in \mathbb{R}^n$. We then get $p(x) = \sum_{j=1}^k (c_j^T x)^2 = x^T (\sum_{j=1}^k c_j c_j^T) x = x^T C^T C x,$

where $C \in \mathbb{R}^{k \times n}$ is a matrix containing the vectors $c_j$ as its rows.

Thus the two conditions are equivalent.

$n = 2$. A homogeneous polynomial of degree $n$ in two variables $x, y$ can be written as $p(x, y) = \sum_{k=0}^d c_k x^k y^{d-k}$. We suppose that the polynomial is not identically zero (in which case it is trivially in both $P_{d,2}$ and $\Sigma_{d,2}$). Then we may suppose, by making a linear change of coordinates if necessary, that $c_d \neq 0$. Let $z_1, \ldots, z_d$ be the roots of the polynomial $\sum_{k=0}^d c_k z^k$. Then this polynomial factorizes as $c_d \prod_{k=1}^d (z - z_k)$. Accordingly, we obtain

$$p(x, y) = c_d \prod_{k=1}^d (x - z_k y).$$

Suppose now that $p \in P_{d,2}$. Then $c_d = p(1, 0) > 0$ and can be written as a square $c_d = (\sqrt{c_d})^2$. The multiplicity of every real root $z_k$ must be even, otherwise $p$ becomes negative near the root $(x, y) = (z_k, 1)$. The corresponding factors in (12) hence group into squares. For every complex root $z_k = a_k + i b_k$ there exists a complex conjugate root $z_{k'} = \bar{z}_k = a_k - i b_k$, and the corresponding product can be written as

$$(x - z_k y)(x - z_{k'} y) = x^2 - 2 a_k x y + (a_k^2 + b_k^2) y^2 = (x - a_k y)^2 + (b_k y)^2.$$ 

The polynomial (12) is then a sum of squares, which shows $p \in \Sigma_{d,2}$.

Example: The Motzkin polynomial $p(x, y, z) = x^4 y^2 + x^2 y^4 + z^6 - 3 x^2 y^2 z^2$ is in $P_{6,3}$ by the arithmetic-geometric inequality, but not in $\Sigma_{6,3}$.

That not every nonnegative polynomial can be represented as a sum of squares of polynomials with lower degree was already known to Hilbert in the 19th century. At the 2nd ICM in 1900 he posed the following question:

Hilberts 17th problem: Can every nonnegative polynomial be represented as a sum of squares of rational functions?

The question was positively answered by Artin in the 30s.

Consider a real symmetric $n \times n$ matrix $A$ and the quartic polynomial

$$p(x) = \sum_{i,j=1}^n A_{ij} x_i^2 x_j^2$$

on $\mathbb{R}^n$. We have $p \in P_{4,n}$ if and only if the quadratic polynomial $\sum_{i,j=1}^n A_{ij} x_i x_j = x^T A x$ takes nonnegative values for all $x \in \mathbb{R}^n$, i.e., if the matrix $A$ is copositive. This is NP-hard to decide, however.

To detect whether a given polynomial is in the cone $P_{2d,n}$ is therefore in general a difficult problem. In contrast to this stands the easy algorithmic accessibility of the cone of sums of squares $\Sigma_{2d,n}$.

Let us devise an algorithm to check whether a given polynomial $p$ is an element of $\Sigma_{2d,n}$. To this end, form the vector $x$ of monomials $\prod_{k=1}^n x_k^{\alpha_k}$ of degree $d$ in the variables $x_k$. The exponents $\alpha_k$ are hence nonnegative integers which sum to $d$. Let $N$ be the size of the vector $x$.

Suppose there exists a positive semi-definite matrix $A \in S^N_+$ such that $p(x) = x^T A x$. Factor the matrix $A$ as $A = B^T B$ with $B \in \mathbb{R}^{k \times N}$, and let $b_j$ be the rows of $B$. Then we obtain

$$p(x) = x^T B^T B x = \sum_{j=1}^k (b_j, x)^2,$$
and $p$ has been represented as a sum of squares of $k$ polynomials $q_j(x) = \langle b_j, x \rangle$ of degree $d$. Thus $p \in \Sigma_{2d,n}$.

On the other hand, suppose that $p \in \Sigma_{2d,n}$. Then there exist $k$ homogeneous polynomials $q_1(x), \ldots, q_k(x)$ of degree $d$ such that $p(x) = \sum_{j=1}^{k} q_j^2(x)$. Every polynomial $q_j(x)$ can be written as a scalar product $\langle c_j, x \rangle$ for some vector $c_j \in \mathbb{R}^N$. Let $C \in \mathbb{R}^{k \times N}$ be the matrix whose rows are the vectors $c_j$. Then we get

$$p(x) = \sum_{j=1}^{k} \langle c_j, x \rangle^2 = x^T C^T C x,$$

and the polynomial $p$ has been written as $x^T A x$ with $A$ positive semi-definite.

We obtain the following result.

**Lemma 11.4.** A homogeneous polynomial $p$ of degree $2d$ in $n$ variables $x_1, \ldots, x_n$ is an element of the cone $\Sigma_{2d,n}$ if and only if there exists a positive semi-definite real symmetric $N \times N$ matrix $A$ such that $p(x) = x^T A x$.

Apart from the conic constraint $A \in S_n^+$, this imposes a finite number of equality relations on $A$ which are jointly linear in the coefficients of $p$ and the elements of $A$. The existence of such a matrix $A$ can hence be incorporated as a constraint into a semi-definite program involving the coefficients of the polynomial $p$.

**Example:** We want to check whether $p(x, y) = \sum_{j=0}^{2d} c_j x^j y^{2d-j} \in \Sigma_{2d,2}$. Let us form the vector $x = (x^d, x^{d-1} y, \ldots, y^d)^T$ of length $N = d + 1$. We shall index the elements of $x$ as well as the elements of $N \times N$ matrices from 0 to $d$ for convenience. Then we have $x^T A x = \sum_{i,j=0}^{d} A_{ij} x_i x_j = \sum_{i,j=0}^{n} A_{ij} x_i (d-i) + (d-j) y^{i+j} = \sum_{k=0}^{2d} x^{2d-k} y^k \sum_{i,j;i+j=k} A_{ij}$. The condition $p(x) = x^T A x$ can then be written as

$$\sum_{i,j;i+j=k} A_{ij} = c_k, \quad \forall k = 0, \ldots, 2d.$$  

(13)

In other words, the sums of the elements of $A$ on the skew-diagonals have to equal the coefficients of the polynomial $p$. We get the following result.

**Lemma 11.5.** A polynomial $p(x, y) = \sum_{j=0}^{2d} c_j x^j y^{2d-j}$ is nonnegative if and only if there exists a positive semi-definite matrix $A \in S_{d+1}^+$ such that (13) holds.

**Example:** We want to check whether the polynomial

$$p(x, y, z) = c_{400} x^4 + c_{310} x^3 y + \cdots + c_{013} y z^3 + c_{004} z^4$$

is in $\Sigma_{4,3}$. Let us form the vector $x = (x^2, y^2, z^2, yz, xz, xy)^T$ of length $N = 6$. Then we have

$$x^T A x = A_{11} x^4 + A_{22} y^4 + A_{33} z^4 + (2A_{12} + A_{66}) x^2 y^2 + (2A_{13} + A_{55}) x^2 z^2 + (2A_{23} + A_{44}) y^2 z^2 + (2A_{14} + 2A_{56}) x^2 y z + (2A_{25} + 2A_{46}) x y^2 z + (2A_{36} + 2A_{45}) x y z^2 + 2A_{15} x^3 z + 2A_{16} x^2 y + 2A_{24} y^3 z + 2A_{34} y^2 z^2 + 2A_{35} x z^3.$$

The condition $p(x) = x^T A x$ can then be written as

$$
\begin{align*}
A_{11} &= c_{400}, \\
A_{22} &= c_{404}, \\
A_{33} &= c_{004}, \\
2A_{12} + A_{66} &= c_{220}, \\
&\quad \vdots \\
2A_{34} &= c_{013}, \\
2A_{35} &= c_{103}.
\end{align*}
$$

The existence of a positive semi-definite matrix $A \in S_{d+1}^+$ satisfying these linear equations is then equivalent to the inclusion $p \in \Sigma_{4,3}$ and hitherto to the nonnegativity of the polynomial $p(x, y, z)$.  

For some matrix \( B \) exist and this matrix is of the form \( A \). The diagonal elements of the block \( K \) are zero, and this matrix is of the form \( A \). Thus \( K \) depend only on the elements of the block \( A \) and the diagonal elements of the block \( A^2 \). We can therefore assume that all other elements of the matrix \( A \) are zero, and this matrix is of the form
\[
A = \text{diag}(B, c_{12}, c_{13}, \ldots, c_{n-1,n})
\]
for some matrix \( B \in S^n_+ \) and some nonnegative scalars \( c_{ij} \), \( 1 \leq i < j \leq n \). We get
\[
\mathbf{x}^T A \mathbf{x} = \sum_{i,j=1}^{n} B_{ij} x_i^2 x_j^2 + \sum_{i<j} c_{ij} x_i^2 x_j^2.
\]
Comparing coefficients with \( p_A(x) = \sum_{i,j=1}^{n} A_{ij} x_i^2 x_j^2 \) we obtain that the matrix \( A \) is in \( K_0 \) if and only if there exist \( B \in S^n_+ \) and \( c_{ij} \geq 0 \), \( 1 \leq i < j \leq n \), such that \( \lim A = \text{diag} B \) and \( A_{ij} = B_{ij} + c_{ij} \) for all \( i < j \). Thus \( K_0 = S^n_+ + N^n \), where \( N^n \) is the cone of element-wise nonnegative matrices with zero diagonal (it is easily seen that this last condition can be dropped).

Diananda proved the following result in 1962:

**Theorem 11.6.** The equality \( C^n = S^n_+ + N^n \) holds if and only if \( n \leq 4 \).

We can strengthen the inner approximation \( K_0 \) by defining the following hierarchy of cones, parameterized by an integer \( r \geq 0 \).
\[
K_r = \left\{ A \in S^n \mid \left( \sum_{j=1}^{n} x_j^2 \right)^r p_A(x) \in \Sigma_{4+2r,n} \right\}.
\]
We have the following result:

**Theorem 11.7.** Let \( A \in \text{int} C^n \). Then there exists an \( r \geq 0 \) such that \( A \in K_{r'} \) for all \( r' \geq r \).

The approximations of \( C^n \) by \( K_r \) are increasingly tight, but become also more complex.

### 11.2 Sums of squares relaxations for polynomial optimization problems

Let us now consider how the approximation of the cone \( P_{d,n} \) of nonnegative polynomials by the cone \( \Sigma_{d,n} \) of sums of squares polynomials allows to approximate difficult optimization problems with polynomial data by easily solvable semi-definite programs.

**Definition 11.8.** A set \( K \subset \mathbb{R}^n \) is called **basic semi-algebraic** if it is of the form
\[
K = \{ x \mid f_i(x) = 0, \quad g_j(x) \leq 0 \}
\]
for some polynomials \( f_i, g_j : \mathbb{R}^n \to \mathbb{R} \).

A set \( K \subset \mathbb{R}^n \) is called **semi-algebraic** if it is a union of a finite number of basic semi-algebraic sets.

Our general optimization problem will be to find the minimum of a polynomial on a semi-algebraic set \( K \):
\[
\min_{x \in K} f_0(x).
\]
Minimizing over a union \( K = \bigcup_{j} K_j \) of sets is equivalent to minimizing over each set \( K_j \) and taking the minimum of the results. Therefore we may assume without loss of generality that \( K \) is already a basic semi-algebraic set given by (14).
In order to reformulate the problem we introduce the cone $P_{d,K}$ of polynomials of degree not exceeding $n$ which are nonnegative on the basic semi-algebraic set $K$. This is a finite-dimensional closed convex cone. Then we may write above problem as

$$\max \tau : \quad f_0(x) - \tau \in P_{d,K},$$

where $d$ is not smaller than the degree of $f_0$.

The cone $P_{d,K}$ is in general difficult to describe. We replace it by the cone $\Sigma_{d,K}$ consisting of all polynomials $p(x)$ of degree not exceeding $d$ which can be represented as a sum

$$p(x) = \sigma_0(x) + \sum_i p_i(x)f_i(x) - \sum_j \sigma_j(x)g_j(x),$$

where $p_i(x)$ are arbitrary polynomials and $\sigma_0(x), \sigma_j(x)$ are sums of squares of polynomials. Clearly every polynomial in $\Sigma_{d,K}$ is nonnegative on $K$ and hence in $P_{d,K}$, because every term in the above sum is nonnegative on $K$. Moreover, the above decomposition yields equality relations which are jointly linear in the coefficients of $p$ and the unknown polynomials $\sigma_0, p_i, \sigma_j$. Therefore the inclusion $p \in \Sigma_{d,K}$ can be expressed by a finite number of semi-definite conic and linear equality constraints. The cone $\Sigma_{d,K}$ is hence a semi-definite representable inner approximation of $P_{d,K}$, and the approximating problem

$$\max \tau : \quad f_0(x) - \tau \in \Sigma_{d,K}$$

is a semi-definite program.

The relaxation can be strengthened by imposing the condition

$$p(x) \cdot \left( \sum_{i=1}^n x_i^2 \right)^{\tau} = \sum_{j=1}^k (c_j, x)^2 = x^T C^T C x$$

instead of (15).

**Example:** We wish to solve the problem

$$\min x + y : \quad x \geq 0, \ x^2 + y^2 = 1.$$  \hspace{1cm} (16)

The set $K = \{ (x, y) \mid x \geq 0, \ x^2 + y^2 = 1 \}$ is a semi-circle and is already basic semi-algebraic. Choose $d = 3$. We approximate the set $P_{3,K}$ of cubic polynomials which are nonnegative on the semi-circle by the set $\Sigma_{3,K}$ of polynomials which are expressible in the form

$$p(x, y) = \sigma_0(x, y) + l(x, y)(x^2 + y^2 - 1) + \sigma_1(x, y)x,$$

where $\sigma_0, \sigma_1$ are sums of squares of polynomials of degree 2 and $l$ is a linear polynomial. Let us introduce the vector of monomials $x = (x, y, 1)^T$ of degree not exceeding 1. Then $p \in \Sigma_{3,K}$ if and only if $p$ can be written as

$$p(x, y) = x^T A^0 x + 1^T x \cdot (x^2 + y^2 - 1) + (x^T A^1 x) \cdot x$$

$$= (A_{11}^0 + l_x)x^3 + (2A_{12}^1 + l_y)x^2y + (A_{11}^0 + 2A_{13}^1 + l_1)x^2 + (A_{12}^1 + l_x)xy^2 + (2A_{12}^0 + 2A_{23}^1)xy$$

$$+ (2A_{03}^0 + A_{33}^1 - l_x)x + l_yy^3 + (A_{12}^0 + l_1)y^2 + (2A_{23}^0 - l_y)y + A_{33}^0 - l_1,$$

where $l = (l_x, l_y, l_1)^T \in \mathbb{R}^3$ and $A^0, A^1 \in S_3^+$.

Hence the semi-definite program approximating the original problem can be written as

$$\max_{A^0, A^1 \in S_3^+} \tau : \quad A_{11}^1 + l_x = 2A_{12}^0 + l_y = A_{11}^0 + 2A_{13}^1 + l_1 = A_{22}^1 + l_x = 2A_{12}^0 + 2A_{23}^1 = l_y = A_{22}^0 + l_1 = 0,$$

$$2A_{13}^0 + A_{33}^1 - l_x = 2A_{23}^0 - l_y = 1, \ A_{33}^0 - l_1 = -\tau.$$  \hspace{1cm}

Using the linear equalities to eliminate variables this leads to the equivalent SDP

$$\max -(A_{33}^0 + A_{11}^1 + 2A_{13}^1) : \quad \begin{pmatrix} A_{11}^1 & A_{12}^0 & A_{13}^0 \\ A_{12}^0 & A_{11}^1 + 2A_{13}^1 & \frac{1}{2} \\ A_{13}^0 & \frac{1}{2} & A_{33}^0 \end{pmatrix} \succeq 0, \quad \begin{pmatrix} A_{11}^1 & 0 & A_{13}^0 \\ 0 & A_{11}^1 & -A_{12}^0 \\ A_{13}^0 & -A_{12}^0 & 1 - A_{11}^1 - 2A_{13}^0 \end{pmatrix} \succeq 0.$$  \hspace{1cm}

Its solution yields the optimal value $-1$. 

11.3 Moment relaxations

Let \( \mu \) be a nonnegative measure on \( \mathbb{R}^n \) with support \( \text{supp} \mu \). The set of nonnegative measures with support in some set \( K \subset \mathbb{R}^n \) forms a convex cone. If \( K \) consists of more than a finite number of points, then this cone has infinite dimension. The extremal measures in this cone are given by the multiples of the \( \delta \)-functions \( \mu(x) = \delta(x - \hat{x}) \), where \( \hat{x} \in K \). The measure \( \delta(x - \hat{x}) \) has support \( \{ \hat{x} \} \) and evaluates on functions as

\[
\int_{\mathbb{R}^n} f(x) \delta(x - \hat{x}) \, dx = f(\hat{x}).
\]

Let \( \mathbb{R}^n \) be indexed by the coordinates \( x_1, \ldots, x_n \).

**Definition 11.9.** Let \( \alpha = (\alpha_1, \ldots, \alpha_n) \) be a vector of nonnegative integers. The moment \( m_\alpha \) of the measure \( \mu \) is the value of the integral

\[
m_\alpha(\mu) = \int_{\mathbb{R}^n} x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \cdots \cdot x_n^{\alpha_n} \mu(x) \, dx = \int_{\mathbb{R}^n} x^\alpha \mu(x) \, dx.
\]

Here and in the sequel we shall use the notation \( x^\alpha \) for the product \( \prod_{i=1}^n x_i^{\alpha_i} \).

A given moment \( m_\alpha \) is a linear functional on the cone of measures. Not all moments may exist for a given measure, because the integral may diverge. For the \( \delta \)-function \( \mu(x) = \delta(x - \hat{x}) \) all moments exist, however, and are given by \( m_\alpha(\mu) = \delta^\alpha \).

Since we work only with finite-dimensional objects, we shall fix a degree \( d \) and consider only moments \( m_\alpha \) for which \( |\alpha| = \sum_{i=1}^n \alpha_i \) does not exceed \( d \). The set of such index vectors \( \alpha \) has finite cardinality \( N \) and gives rise to an \( N \)-dimensional moment vector \( m(\mu) = (m_\alpha(\mu))_{|\alpha| \leq d} \).

The moment cone \( M_d \subset \mathbb{R}^N \) is then the set of all vectors which can be produced as moment vectors of some nonnegative measure \( \mu \). For subsets \( K \subset \mathbb{R}^n \), we shall also consider the cones \( M_{d,K} \subset \mathbb{R}^N \) consisting of moment vectors of measures \( \mu \) with support in \( K \). The moment cones can be seen as finite-dimensional projections of the infinite-dimensional cone of measures.

The moment cones \( M_{d,K} \) are in general difficult to describe. We shall consider necessary conditions which a moment vector \( m(\mu) \) of a nonnegative measure has to satisfy. The set of vectors satisfying these conditions will then yield an *outer* approximation of the moment cone.

Let \( x = (1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^{\lfloor d/2 \rfloor})^T \) be the vector of monomials \( x^\alpha \) for \( |\alpha| \) not exceeding the integer part of \( d/2 \). Then all entries in the rank 1 matrix \( xx^T \) will be monomials of degree not exceeding \( d \). Consider the matrix-valued integral

\[
\int_{\mathbb{R}^n} xx^T \mu(x) \, dx.
\]

This is a positive semi-definite matrix whose entries are elements of the moment vector \( m(\mu) \). We therefore obtain a semi-definite conic constraint on the moment vector, namely that the above matrix should be in the cone of positive semi-definite matrices.

Let now \( K = \{ x \in \mathbb{R}^n \mid f_i(x) = 0, \ g_j(x) \leq 0 \} \) be a basic semi-algebraic set, and let \( \mu \) be a measure with support in \( K \).

Let \( d_i \) be the degree of the polynomial \( f_i \). Then for every polynomial \( p \) of degree not exceeding \( d - d_i \) we have

\[
\int_{\mathbb{R}^n} p(x) f_i(x) \mu(x) \, dx = 0.
\]

On the other hand, the left-hand side is a linear combination of elements of the moment vector \( m(\mu) \). This yields a linear equality relation on the moment vector \( m(\mu) \). A maximal linearly independent set of such equalities can be obtained if \( p(x) \) runs through all monomials \( x^\beta \) with \( |\beta| \leq d - d_i \).

Let now \( d_j \) be the degree of the polynomial \( g_j \) and let \( q(x) \) be a polynomial which is nonnegative on \( K \). Then we obtain

\[
\int_{\mathbb{R}^n} q(x) g_j(x) \mu(x) \, dx \leq 0.
\]

This leads in a similar way to a linear inequality relation on \( m(\mu) \).
We may also form the vector \( x' \) of all monomials with degree not exceeding the integer part of \( \frac{d-d_j}{2} \) and consider the matrix-valued integral

\[
- \int_{\mathbb{R}^n} x'(x')^T g_j(x) \mu(x) \, dx.
\]

This integral evaluates to a positive semi-definite matrix and every of its entries is a linear combination of elements of \( m(\mu) \). This yields a semi-definite conic constraint on \( m(\mu) \).

Let us now consider the problem

\[
\min_{x \in K} f_0(x),
\]

where \( f_0 = \sum c_\alpha x^\alpha \) is a polynomial of degree not exceeding some integer \( d \), and \( K \) is a basic semi-algebraic set as above. We can rewrite this problem equivalently as

\[
\min_{\mu : \text{supp } \mu \subset K} \int_{\mathbb{R}^n} f_0(x) \mu(x) \, dx : \int_{\mathbb{R}^n} \mu(x) \, dx = 1.
\]

Here the minimization is performed over all probability measures with support in \( K \).

The equality condition on \( \mu \) can, however, be written as \( m_0(\mu) = 1 \), and the integral in the cost function evaluates to the linear combination \( \sum c_\alpha m_\alpha(\mu) \) of elements of the moment vector \( m(\mu) \). The problem thus becomes

\[
\min_{m \in M_{d,K}} \sum c_\alpha m_\alpha : \quad m_0 = 1.
\]

Replacing the difficult condition \( m \in M_{d,K} \) by a set of semi-definite and linear constraints like those constructed above then yields a semi-definite approximation of the problem.

Example: Let us again consider problem (16). Set \( d = 3 \), then the moment vector is 10-dimensional. We obtain the SDP

\[
\min m_{10} + m_{01} : \begin{pmatrix}
  m_{00} & m_{10} & m_{01} \\
  m_{10} & m_{20} & m_{11} \\
  m_{01} & m_{11} & m_{02}
\end{pmatrix} \succeq 0, \quad m_{20} + m_{02} - m_{00} = m_{30} + m_{12} - m_{10} = m_{21} + m_{03} - m_{01} = 0,
\]

\[
\begin{pmatrix}
  m_{10} & m_{20} & m_{11} \\
  m_{20} & m_{30} & m_{21} \\
  m_{11} & m_{21} & m_{12}
\end{pmatrix} \succeq 0, \quad m_{00} = 1.
\]

Its solution also yields the optimal value \(-1\).

References