where *"sdp max-cut"* is the optimal value obtained by solving (7).

The following randomized algorithm proposed by Goemans and Williamson in 1995[3] gives an extraorinary bound for the max-cut problem.

In (7), it is optimal solution $\boldsymbol{X}$ belongs to $\mathcal{S}^n_+$, and therefore, it is a Gram matrix and $\exists \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^\ell$ such that

$$X_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle \qquad (i, j = 1, 2, \ldots, n).$$

Moreover, since $X_{ii} = 1$, $\|\boldsymbol{v}_i\| = 1$. Such n vectors can be obtained using the Cholesky decomposition for instance. Once we have determined $\boldsymbol{v}_i \quad (i = 1, 2, \ldots, n)$, we execute the following random algorithm.

Set maxcut:$= -\infty$.
For $k := 1$ **to** MAX
  Choose a vector $\boldsymbol{v} \in \mathbb{R}^\ell$ uniformly distributed in $S^{\ell-1} := \{\boldsymbol{x} \in \mathbb{R}^\ell \mid \|x\|_2 = 1\}$.
  Define a cut $S_k \subseteq V$ consisting of i with $\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle \geq 0$.
  Compute $\delta(S_k)$.
  If $\delta(S_k) >$ maxcut,
    then maxcut:$= \delta(S_k)$.

Theorem 4.1 (Goemans-Williamson (1995)) The above algorithm provides the following expectation bound:

$$E[\text{rand SDP}] \geq 0.8785[\text{opt. max-cut}].$$

Proof:

First, let us compute the probability of an edge $(i, j) \subseteq E$ being selected by the above procedure.

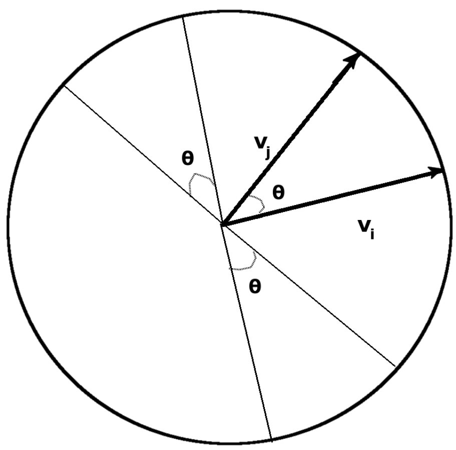

Figure 2: Plane defined by the vectors $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, the angle θ between them, and the slice of $S^{\ell-1}$.

Figure 2 shows the plane defined by the vectors $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, and also the slice of $S^{\ell-1}$. We can see from Figure 2, that since $\|\boldsymbol{v}_i\| = \|\boldsymbol{v}_j\| = 1$, the probability of $\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle$ and $\langle \boldsymbol{v}, \boldsymbol{v}_j \rangle$ have opposite signs

[3] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maxicum cut and satisfiability problems using semidefinite programming," *J. Assoc. Comput. Mach.*, **42** (1995), pp. 1115–1145.

is $\frac{\arccos(X_{ij})}{\pi}$. Recall that $X_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \|\boldsymbol{v}_i\|_2 \|\boldsymbol{v}_j\|_2 \cos\theta$. Also $|X_{ij}| \leq 1 \quad (i,j = 1,2,\ldots,n)$ by Exercise 2.

Therefore, the expectation of the capacity which can be obtained by the algorithm is:

$$E[\text{rand SDP}] = \sum_{i,j=1}^{n} \frac{w_{ij}}{2} \frac{\arccos(X_{ij})}{\pi} \geq \sum_{i,j=1}^{n} \frac{w_{ij}}{2} \frac{\alpha}{2}(1-X_{ij}) = \alpha \sum_{i,j=1}^{n} \frac{w_{ij}}{4}(1-X_{ij}) \geq \alpha[\text{opt. max-cut}],$$

for $\alpha = 0.8785$.

Now, it remains to show that

$$\frac{\arccos(x)}{\pi} \geq \frac{\alpha}{2}(1-x) \qquad \forall x \in [-1,1].$$

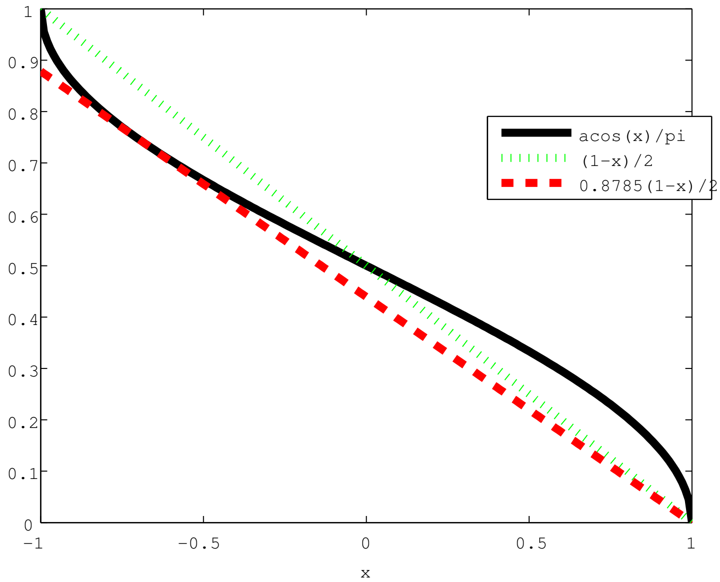

Figure 3: Function values for $\frac{\arccos(x)}{\pi}$, $\frac{\alpha}{2}(1-x)$, and $\frac{(1-x)}{2}$ for $x \in [-1,1]$.

This can be seen if we plot their values for $x \in [-1,1]$ as in Figure 3. ∎

It is reported that actual numerical experiments give an approximation better than 0.9 of the optimal value. On the other hand, it is also known that the theoretical expectation can not be better than the bound $16/17 \approx 0.9412$. [4]

4.4 Extension to the Maximization of a Convex Quadratic Function

The ideia of Goemans-Williamson approach can be extended to the following problem.

$$\begin{cases} \text{maximize} & \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in \{-1,1\}^n \end{cases} \tag{8}$$

We can assume without loss of generality that the matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is symmetric, positive semidefinite, and diagonally dominant, *i.e.*, $Q_{ii} \geq \sum_{j\neq i}^{n} |Q_{ij}|$. See Exercise 1.

Again, a semidefinite program relaxation of it will be:

[4] J. Håstad, "Some optimal inapproximability results," *J. Assoc. Comput. Mach.*, **48** (2001), pp. 798–859.

$$\begin{cases} \text{maximize} & \langle \boldsymbol{Q}, \boldsymbol{X} \rangle \\ \text{subject to} & X_{ii} = 1, \quad i = 1, 2, \ldots, n \\ & \boldsymbol{X} \in \mathcal{S}_+^n \end{cases} \tag{9}$$

The following result is given by Nesterov[5].

Theorem 4.2 (Nesterov) For $\boldsymbol{Q} \in \mathcal{S}_+^n$ in (8), we have

$$[\text{sdp qp}] \geq [\text{opt. qp}] \geq \frac{2}{\pi}[\text{sdp qp}] \qquad (\frac{2}{\pi} = 0.6366...)$$

where *"sdp qp."* is the optimal value of (9) and *"opt. qp."* is the optimal value of (8).

Proof:

The first inequality is obvious because (9) is an SDP relaxation of (8). Similar to the proof of Goemans-Williamson's result, let $\boldsymbol{X} \in \mathcal{S}_+^n$ be any feasible solution of (9) and $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^\ell$ such that $X_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle$. We chose a vector $\boldsymbol{v}$ uniformily distributed in $S^{\ell-1}$, and define a vector $\boldsymbol{x} \in \mathbb{R}^n$ by the following process. Its elements will be equal to $\text{sign}(\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle)$ for $i = 1, 2, \ldots, n$. It is clear that $\boldsymbol{x}$ is feasible for (8). The expectation of the objective function calculated for this random variable is:

$$o := \sum_{i,j=1}^n Q_{ij} E_{\boldsymbol{v}}[\text{sign}(\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle)\text{sign}(\langle \boldsymbol{v}, \boldsymbol{v}_j \rangle)].$$

The probability of x_i and x_j have the same sign is $\frac{\pi - \arccos(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)}{\pi}$ and opposite signs $\frac{\arccos(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)}{\pi}$ (see Figure 2). Therefore,

$$\begin{aligned} o &= \sum_{i,j=1}^n Q_{ij} \left(\frac{\pi - \arccos(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)}{\pi} - \frac{\arccos(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)}{\pi} \right) \\ &= \sum_{i,j=1}^n Q_{ij} \frac{2}{\pi} \left(\frac{\pi}{2} - \arccos(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle) \right) \\ &= \sum_{i,j=1}^n Q_{ij} \frac{2}{\pi} \arcsin(\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle) = \frac{2}{\pi} \sum_{i,j=1}^n Q_{ij} \arcsin(X_{ij}). \end{aligned}$$

Since this value is just the expected value, and $\boldsymbol{X}$ is any feasible solution of (9), we have in fact that

$$[\text{opt. qp}] \geq \frac{2}{\pi} \max\{\langle \boldsymbol{Q}, \arcsin(\boldsymbol{X}) \rangle \mid \boldsymbol{X} \in \mathcal{S}_+^n, \ X_{ii} = 1 \ (i = 1, 2, \ldots, n)\},$$

where $\arcsin(\boldsymbol{X})$ is the matrix with elements equal to $\arcsin(X_{ij})$. Finally, since $\boldsymbol{Q} \in \mathcal{S}_+^n$, $X_{ii} = 1$ and $\boldsymbol{X} \in \mathcal{S}_+^n$, by Lemma 4.3, $\langle \boldsymbol{Q}, \arcsin(\boldsymbol{X}) - \boldsymbol{X} \rangle \geq 0$, and therefore

$$[\text{opt. qp}] \geq \frac{2}{\pi} \max\{\langle \boldsymbol{Q}, \boldsymbol{X} \rangle \mid \boldsymbol{X} \in \mathcal{S}_+^n, \ X_{ii} = 1 \ (i = 1, 2, \ldots, n)\} = \frac{2}{\pi}[\text{sdp qp}].$$

∎

Lemma 4.3 Let $\boldsymbol{X} \in \mathcal{S}_+^n$ with diagonal elements equal to one. Then,

$$\arcsin(\boldsymbol{X}) - \boldsymbol{X} \in \mathcal{S}_+^n.$$

[5] Yu. Nesterov, "Quality of semidefinite relaxation for nonconvex quadratic optimization," *CORE Discussion Paper*, 1997.

Proof: Since the diagonal elements of $\boldsymbol{X}$ are all ones, by Exercise 2, $|X_{ij}| \leq 1 \quad (i,j = 1,2,\ldots,n)$. Then the following Taylor expansion converges for all elements of $\boldsymbol{X}$ which are in $[-1,1]$.

$$\arcsin(\boldsymbol{X}) - \boldsymbol{X} = \sum_{k=1}^{\infty} \frac{1\cdot 3\cdot 5\cdot (2k-1)}{2^k k!(2k+1)} \boldsymbol{X}^{2k+1},$$

where $\boldsymbol{X}^k$ denotes the matrix with the elements equal to X_{ij}^k. Since the Hadamar product of positive semidefinite matrices is positive semidefinite, the right hand side of the above equation is positive semidefinite and the result follows. ∎

4.5 Exercises

1. Show that for problem (8), one can always assume that $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is symmetric, positive semidefinite, and diagonally dominant, *i.e.*, $Q_{ii} \geq \sum_{j\neq i}^{n} |Q_{ij}|$.
2. Show that for $\boldsymbol{X} \in \mathcal{S}_+^n$ and all diagonal elements equal to one we have $|X_{ij}| \leq 1 \quad (i,j = 1,2,\ldots,n)$.

5 Polynomial Optimization Problem (POP)

Definition 5.1 Let $\mathbb{R}[\boldsymbol{x}]$ denote the ring of real polynomials in the variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. A polynomial $p \in \mathbb{R}[\boldsymbol{x}]$ is a *sum of squares* (s.o.s.) if p can be written in the following form:

$$p(\boldsymbol{x}) = \sum_{j\in J} p_j(\boldsymbol{x})^2, \quad \boldsymbol{x} \in \mathbb{R}^n$$

for some finite family of polynomials $\{p_j \mid j \in J\} \subseteq \mathbb{R}[\boldsymbol{x}]$. Therefore, the degree of p is even and the maximum degree of p_j is half of it. We denote the set of all s.o.s. polynomials by $\Sigma[\boldsymbol{x}] \subseteq \mathbb{R}[\boldsymbol{x}]$.

For a multi-index $\boldsymbol{\alpha} \in \mathbb{N}^n$, let $|\boldsymbol{\alpha}| := \sum_{i=1}^n \alpha_i$. Denote by

$$\boldsymbol{v}_d(\boldsymbol{x}) := (\boldsymbol{x}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}|\leq d} := (1, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^2, \ldots, x_1^d, \ldots, x_n^d)^T$$

the vector formed by all monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ of degree less or equal to d. This vector has dimension $s(d) := \begin{pmatrix} n+d \\ d \end{pmatrix}$. Those monomials also form the canonical basis of the vector space $\mathbb{R}[\boldsymbol{x}]_d$ of polynomials of degree less or equal to d.

Proposition 5.2 A polynomial $g \in \mathbb{R}[\boldsymbol{x}]_{2d}$ has a s.o.s. decomposition if and only if there exists $\boldsymbol{Q} \in \mathcal{S}_+^{s(d)}$ such that $g(\boldsymbol{x}) = \boldsymbol{v}_d(\boldsymbol{x})^T \boldsymbol{Q} \boldsymbol{v}_d(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Proof:
g has a s.o.s decomposition $\Leftrightarrow$ $\exists k \in \mathbb{N}$ and $\exists \boldsymbol{U} = (\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots \boldsymbol{u}_k) \in \mathbb{R}^{s(d)\times k}$ such that $\boldsymbol{g} = \sum_{i=1}^k (\boldsymbol{u}_i^T \boldsymbol{v}_d(\boldsymbol{x}))^2 = \boldsymbol{v}_d(\boldsymbol{x})^T \boldsymbol{Q} \boldsymbol{v}_d(\boldsymbol{x})$ for $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{U}^T$.
Now, $\boldsymbol{Q} \in \mathcal{S}_+^{s(d)}$ is a Gram matrix $\Leftrightarrow$ $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{U}^T$ for some $\boldsymbol{U} \in \mathbb{R}^{s(d)\times k}$. ∎

Example 5.3 $g(x_1,x_2) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4$ has a s.o.s. decomposition since

$$\begin{aligned}
g(x_1,x_2) &= \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix} \\
&= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 2 & 0 \\ -3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix} \\
&= \frac{1}{2}(2x_1^2 - 3x_2^2 + x_1x_2)^2 + \frac{1}{2}(x_2^2 + 3x_1x_2)^2.
\end{aligned}$$