

Now, let us turn on to the problem of finding the *maximum cut*. It is known that this problem is *NP-complete*, i.e., it is *Non-deterministic Polynomial time (NP)* and *Non-deterministic Polynomial time hard (NP-hard)*.

Suppose we have a cut $S \subseteq V := \{1, 2, \dots, n\}$, and we assign the value 1 for x_i if $i \in S$ and -1 for x_i if $i \in V \setminus S$. Also, let us assume that $w_{ij} := 0$ for $(i, j) \notin E$. Then, the capacity of a cut S will be

$$\frac{1}{4} \sum_{i,j=1}^n w_{ij}(1 - x_i x_j), \quad x_i = \begin{cases} 1, & i \in S \\ -1, & i \in V \setminus S \end{cases}$$

Therefore, we can formulate the max-cut problem as an optimization problem.

$$\begin{cases} \text{maximize} & \frac{1}{4} \sum_{i,j=1}^n w_{ij}(1 - x_i x_j) \\ \text{subject to} & x_i^2 = 1 \quad i = 1, 2, \dots, n. \end{cases} \quad (6)$$

For a *maximization* problem, we say that one algorithm gives an α -*approximation* if for an $\alpha \in (0, 1)$, we have

$$[\text{approximate value obtained by the algorithm}] \geq \alpha [\text{optimal value of the problem}]$$

for all instance of the problem class.

4.2.1 A Very Simple Randomized Algorithm

Let us consider the most simple randomized algorithm for the max-cut problem.

```

Set maxcut :=  $-\infty$ .
For  $k := 1$  to MAX
  Choose a cut  $S_k \subseteq V$  such that each vertex of  $V$  is chosen with probability  $\frac{1}{2}$ .
  Compute  $\delta_w(S_k)$ .
  If  $\delta_w(S_k) > \text{maxcut}$ ,
    then maxcut :=  $\delta_w(S_k)$ .

```

Since each possible edge in $G = (V, E)$ can be chosen with probability 0.5, we obtain:

$$E[\text{rand}] = \frac{1}{2} \sum_{1 \leq i \leq j \leq n} w_{ij} = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} \geq \frac{1}{2} [\text{opt. max-cut}].$$

where “*rand*” is the optimal value obtained by the above algorithm and “*opt. max-cut*” is the actual optimal value of the maximum capacity of the cut in the given graph.

4.3 Semidefinite Program Relaxation of the Max-Cut Problem

Associating a new variable X_{ij} for $x_i x_j$ ($1 \leq i \leq j \leq n$), we can rewrite (6) as the following equivalent problem:

$$\begin{cases} \text{maximize} & \frac{1}{4} \langle \mathbf{W}, \mathbf{E} - \mathbf{X} \rangle \\ \text{subject to} & X_{ii} = 1 \quad i = 1, 2, \dots, n, \\ & \mathbf{X} \in \mathcal{S}_+^n, \\ & \text{rank}(\mathbf{X}) = 1, \end{cases}$$

where \mathbf{E} is the matrix with all elements equal to one.

Neglecting the condition $\text{rank}(\mathbf{X}) = 1$, we obtain the following semidefinite program relaxation of the max-cut problem:

$$\begin{cases} \text{maximize} & \frac{1}{4}\langle \mathbf{W}, \mathbf{E} - \mathbf{X} \rangle \\ \text{subject to} & X_{ii} = 1 \quad i = 1, 2, \dots, n, \\ & \mathbf{X} \in \mathcal{S}_+^n. \end{cases} \quad (7)$$

It is clear that

$$[\text{sdp max-cut}] \geq [\text{opt. max-cut}],$$

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

The following randomized algorithm proposed by Goemans and Williamson in 1995³ gives an extraordinary bound for the max-cut problem.

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

$$X_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle \quad (i, j = 1, 2, \dots, n).$$

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

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

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 shows the plane defined by the vectors \mathbf{v}_i and \mathbf{v}_j , and also its slice of $S^{\ell-1}$. We can see from Figure 2, that since $\|\mathbf{v}_i\| = \|\mathbf{v}_j\| = 1$, the probability of $\langle \mathbf{v}, \mathbf{v}_i \rangle$ and $\langle \mathbf{v}, \mathbf{v}_j \rangle$ have opposite signs is $\frac{\arccos(X_{ij})}{\pi}$. Recall that $X_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \|\mathbf{v}_i\|_2 \|\mathbf{v}_j\|_2 \cos \theta$. Also $|X_{ij}| \leq 1$ ($i, j = 1, 2, \dots, n$) by Exercise 3.

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

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

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

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

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

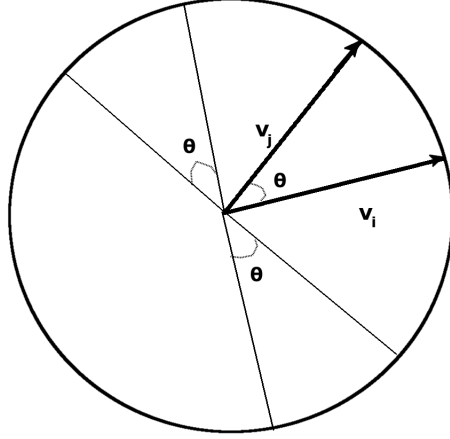


Figure 2: Plane defined by the vectors v_i and v_j , the angle θ between them, and its slice of $S^{\ell-1}$.

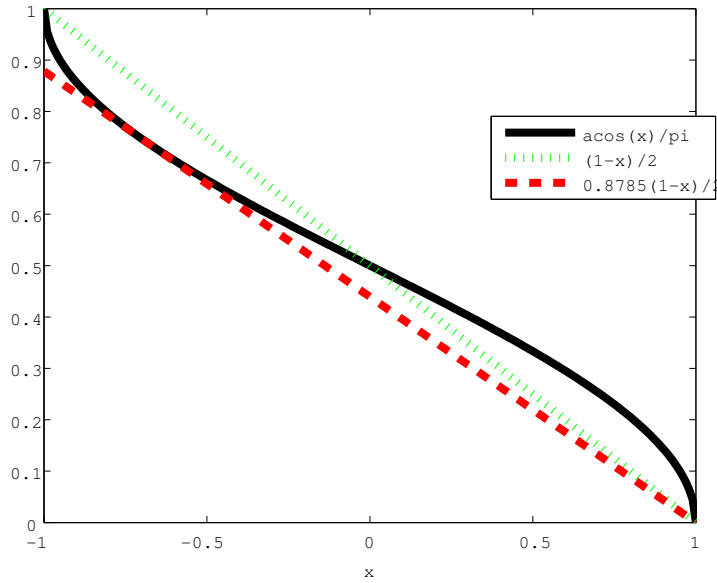


Figure 3: Function values for $\frac{\arccos(x)}{\pi}$, $\frac{\alpha}{2}(1-x)$, and $\frac{(1-x)}{2}$ for $x \in [-1, 1]$.

4.4 Extension to the Maximization of a Convex Quadratic Function

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

$$\begin{cases} \text{maximize} & \mathbf{x}^T \mathbf{Q} \mathbf{x} \\ \text{subject to} & \mathbf{x} \in \{-1, 1\}^n \end{cases} \quad (8)$$

We can assume without loss of generality that the matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is symmetric, positive semidefinite, and diagonally dominant, *i.e.*, $Q_{ii} \geq \sum_{j \neq i}^n |Q_{ij}|$. See Exercise 2.

Again, a semidefinite program relaxation of it will be:

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

The following result is given by Nesterov⁵.

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

$$[\text{sdp qp}] \geq [\text{opt. qp}] \geq \frac{2}{\pi} [\text{sdp qp}] \quad \left(\frac{2}{\pi} = 0.6366\dots \right)$$

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 $\mathbf{X} \in \mathcal{S}_+^n$ be any feasible solution of (9) and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{R}^\ell$ such that $X_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$. We chose a vector \mathbf{v} uniformly distributed in $S^{\ell-1}$, and define a vector $\mathbf{x} \in \mathbb{R}^n$ by the following process. Its elements will be equal to $\text{sign}(\langle \mathbf{v}, \mathbf{v}_i \rangle)$ for $i = 1, 2, \dots, n$. It is clear that \mathbf{x} is feasible for (8). The expectation of the objective function calculated for this random variable is:

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

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

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

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

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

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

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

■

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

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

⁵Yu. Nesterov, “Quality of semidefinite relaxation for nonconvex quadratic optimization,” *CORE Discussion Paper*, 1997.

Proof: Since the diagonal elements of \mathbf{X} are all ones, by Exercise 3, $|X_{ij}| \leq 1$ ($i, j = 1, 2, \dots, n$). Then the following Taylor expansion converges for all elements of \mathbf{X} which are in $[-1, 1]$.

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

where \mathbf{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. \blacksquare

4.5 Exercises

1. Show by constructing an example that a cut corresponding to the maximum cut for a non-negative-weighted graph does not necessary corresponds to the cut for a maximum cut for a graph with positive and negative weights.
2. Show that for problem (8), one can always assume that $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is symmetric, positive semidefinite, and diagonally dominant, *i.e.*, $Q_{ii} \geq \sum_{j \neq i}^n |Q_{ij}|$.
3. Show that for $\mathbf{X} \in \mathcal{S}_+^n$ and all diagonal elements equal to one we have $|X_{ij}| \leq 1$ ($i, j = 1, 2, \dots, n$).

5 Polynomial Optimization Problem (POP)

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

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

for some finite family of polynomials $\{p_j \mid j \in J\} \subseteq \mathbb{R}[\mathbf{x}]$. 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 $\Sigma[\mathbf{x}] \subseteq \mathbb{R}[\mathbf{x}]$.

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

$$\mathbf{v}_d(\mathbf{x}) := (\mathbf{x}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq d} := (1, x_1, \dots, x_n, x_1^2, x_1 x_2, \dots, x_n^2, \dots, x_1^d, \dots, x_n^d)^T$$

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

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

Proof:

g has a s.o.s decomposition $\Leftrightarrow \exists k \in \mathbb{N}$ and $\exists \mathbf{U} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k) \in \mathbb{R}^{s(d) \times k}$ such that $\mathbf{g} = \sum_{i=1}^k (\mathbf{u}_i^T \mathbf{v}_d(\mathbf{x}))^2 = \mathbf{v}_d(\mathbf{x})^T \mathbf{Q} \mathbf{v}_d(\mathbf{x})$ for $\mathbf{Q} = \mathbf{U} \mathbf{U}^T$.

Now, $\mathbf{Q} \in \mathcal{S}_+^{s(d)}$ is a Gram matrix $\Leftrightarrow \mathbf{Q} = \mathbf{U} \mathbf{U}^T$ for some $\mathbf{U} \in \mathbb{R}^{s(d) \times k}$. \blacksquare