Optimización Semidefinida
Clase 08 - Aplicación: max cut

Segundo Cuatrimestre 2021

Facultad de Ciencias Exactas y Naturales, UBA

El problema max cut

Consideramos un grafo no-dirigido $G = (N, E)$, donde N es el conjunto de nodos y $E \subset N \times N$ es el conjunto de aristas. Suponemos $N = \{1, \dots, n\}$.

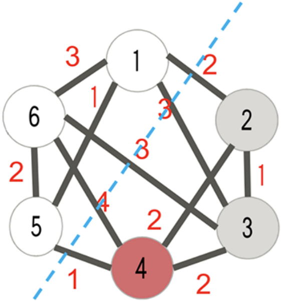
Definimos una matriz de costos no-negativos $W = (w_{ij}) \subset \mathbb{R}_{\geq 0}^{n \times n}$.

Podemos suponer que el grafo es completo, es decir que todos los nodos son adyacentes entre sí, tomando $w_{ij} = 0$ para todas las no-aristas ij; definimos también $w_{ii} = 0$ para todo i.

Como el grafo es no-dirigido, W resulta una matriz simétrica.

Ejemplo

Consideramos el siguiente grafo con los costos indicados.



La matriz de costos es

$$\boldsymbol{W} = \begin{pmatrix} 0 & 2 & 3 & 0 & 1 & 3 \\ 2 & 0 & 1 & 2 & 0 & 0 \\ 3 & 1 & 0 & 2 & 0 & 3 \\ 0 & 2 & 2 & 0 & 1 & 4 \\ 1 & 0 & 0 & 1 & 0 & 2 \\ 3 & 0 & 3 & 4 & 2 & 0 \end{pmatrix}$$

Motivación

Para un subconjunto $K \subset N$ de nodos, definimos el corte $\delta(K)$ como el conjunto de todos las aristas desde K al complemento de K:

$$\delta(K) = \{(i,j) \in E : i \in K, j \notin K\}.$$

Dado un corte K definimos el costo del corte

$$w(\delta(K)) := \sum_{(i,j)\in\delta(K)} w_{ij}.$$

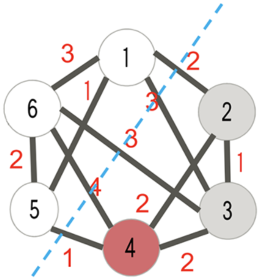
Queremos hallar el corte $\delta(K)$ con el mayor costo posible.

Ejemplo

En el ejemplo anterior, para considerar el corte de la línea celeste, tomamos $K = \{2, 3, 4\}$. El costo del corte $\delta(K)$ es

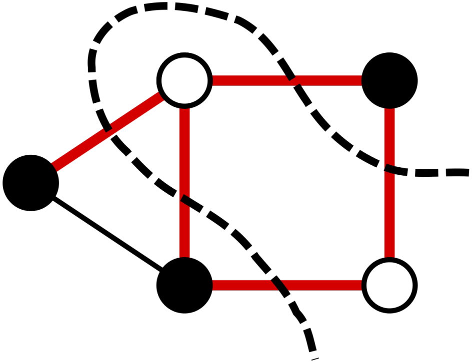
$$w(\delta(K)) = (2) + (3 + 3) + (1 + 4) = 13.$$

Podemos calcularlo sumando los costos de todas las aristas que atraviesa la línea celeste.



Ejemplo

En el siguiente grafo de 5 nodos, asignamos costo 1 a todas las aristas dibujadas y costo 0 a las aristas no dibujadas. Tomando K el conjunto de nodos blancos, obtenemos el corte de mayor costo posible.



Ejemplo

Si consideramos un grafo completo de 4 vértices, con costo 1 en todas las aristas, el costo de un corte queda determinado por la cantidad de nodos en el corte. Obtenemos los siguientes valores:

$$\delta(K) = \begin{cases} 0 & \text{si } \#K = 0, \\ 3 & \text{si } \#K = 1, \\ 4 & \text{si } \#K = 2, \\ 3 & \text{si } \#K = 3, \\ 0 & \text{si } \#K = 4. \end{cases}$$

Por lo tanto, para obtener el corte de mayor costo tomamos un conjunto K de 2 elementos.

Forma matricial

Comenzamos formulando el problema en forma matricial. Queremos calcular el costo de un corte mediante un producto de matrices.

Fijamos un corte K y definimos el vector $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ por

$$x_i = \begin{cases} 1 & \text{si } i \in K \\ -1 & \text{si } i \notin K \end{cases}.$$

De esta forma, $x_i x_j = -1$ si $(i,j) \in \delta(K)$ y $x_i x_j = 1$ si $(i,j) \notin \delta(K)$, y por lo tanto

$$1 - x_i x_j = \begin{cases} 2 & \text{si } (i,j) \in \delta(K), \\ 0 & \text{si } (i,j) \notin \delta(K). \end{cases}$$

Forma matricial

Obtenemos

$$w(\delta(K)) = \sum_{i<j} w_{ij} \frac{1 - x_i x_j}{2} = \frac{1}{4} \sum_{i \neq j} w_{ij}(1 - x_i x_j).$$

Como además $w_{ii} = 0$ para todo i,

$$w(\delta(K)) = \frac{1}{4} \sum_{i} \sum_{j} w_{ij}(1 - x_i x_j) = \left(\frac{1}{4} \sum_{i} \sum_{j} w_{ij} \right) - \frac{1}{4} \sum_{i \neq j} w_{ij}(x_i x_j).$$

Forma matricial

Utilizando $x_i^2 = 1$, reescribimos:

$$\left(\frac{1}{4}\sum_i\sum_j w_{ij}\right) - \sum_{i\neq j} w_{ij}(x_i x_j) = \left(\frac{1}{4}\sum_i\left(\sum_j w_{ij}\right)x_i x_i\right) - \frac{1}{4}\sum_{i\neq j} w_{ij}(x_i x_j).$$

El primer término corresponde a los productos $x_i x_i$ y el segundo a los productos $x_i x_j$, $i \neq j$.

Forma matricial

Definiendo $\boldsymbol{C} \in \mathbb{R}^{n \times n}$ con

$$c_{ij} = -w_{ij}/4 \text{ para } i \neq j$$
$$c_{ii} = \sum_j w_{ij}/4 \text{ para todo } i \ ,$$

obtenemos $w(\delta(K)) = \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x}$.

En el ejemplo, obtenemos $\boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} = 13$ tomando $\boldsymbol{x} = (-1, 1, 1, 1, -1, -1)$,

$$\boldsymbol{W} = \begin{pmatrix} 0 & 2 & 3 & 0 & 1 & 3 \\ 2 & 0 & 1 & 2 & 0 & 0 \\ 3 & 1 & 0 & 2 & 0 & 3 \\ 0 & 2 & 2 & 0 & 1 & 4 \\ 1 & 0 & 0 & 1 & 0 & 2 \\ 3 & 0 & 3 & 4 & 2 & 0 \end{pmatrix} \quad \text{y} \quad \boldsymbol{C} = \frac{1}{4} \begin{pmatrix} 9 & -2 & -3 & 0 & -1 & -3 \\ -2 & 5 & -1 & -2 & 0 & 0 \\ -3 & -1 & 9 & -2 & 0 & -3 \\ 0 & -2 & -2 & 9 & -1 & -4 \\ -1 & 0 & 0 & -1 & 4 & -2 \\ -3 & 0 & -3 & -4 & -2 & 12 \end{pmatrix}.$$

Programación cuadrática

Como cualquier vector $x \in \{-1, 1\}^n$ define un corte, podemos escribir el problema max-cut como el problema de programación cuadrática entera:

$$\begin{aligned} \text{maximizar:} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x}, \\ \text{sujeto a:} \quad & x_i \in \{+1, -1\}, 1 \leq i \leq n. \end{aligned} \tag{IQP}$$

o como un problema cuadrático con restricciones cuadráticas no-convexas

$$\begin{aligned} \text{maximizar:} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x}, \\ \text{sujeto a:} \quad & x_i^2 = 1, 1 \leq i \leq n. \end{aligned} \tag{NQCQP}$$

Ninguna de estas dos formulaciones corresponde a un problema de programación semidefinida. A continuación veremos como relajar las condiciones para obtener un problema SDP.

Relajación - Opción 1

Observamos que (NQCQP) es lineal en los productos $x_i x_j$ y que estos productos son las coordenadas de la matriz $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T \in \mathbb{R}^{n \times n}$ de rango 1.

Más aún, $\boldsymbol{X}$ es simétrica, $x_{ii} = 1$ para todo i y $\boldsymbol{X} \succeq 0$.

Recíprocamente, cualquier matriz de rango 1 con esas propiedades puede escribirse como $\boldsymbol{x}\boldsymbol{x}^T$ para algún vector $\boldsymbol{x} \in \{-1, 1\}^n$. (¿Por qué?)

Finalmente, observamos que

$$\boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} = \sum_{1 \leq i,j \leq n} x_i c_{ij} x_j = \boldsymbol{C} \bullet (\boldsymbol{x}\boldsymbol{x}^T).$$

Relajación - Opción 1

Combinando los resultados, obtenemos que (IQP) es equivalente al problema

$$\begin{array}{ll} \text{maximizar:} & \boldsymbol{C} \bullet \boldsymbol{X} \\ \text{sujeto a:} & x_{ii} = 1, \quad 1 \leq i \leq n \\ & \boldsymbol{X} \succeq 0, \\ & \operatorname{rank}(\boldsymbol{X}) = 1. \end{array}$$

Eliminando la restricción del rango, obtenemos el problema SDP

$$\begin{array}{ll} \text{maximizar:} & \boldsymbol{C} \bullet \boldsymbol{X} \\ \text{sujeto a:} & x_{ii} = 1, \quad 1 \leq i \leq n, \\ & \boldsymbol{X} \succeq 0. \end{array}$$

Vemos que está planteado con un problema en forma primal, excepto que debemos maximizar la función en lugar de minimizarla.

Relajación - Opción 2

Observemos que en (IQP) asociamos a cada nodo $i \in N$ un valor $x_i \in \{-1, 1\}$, que podemos considerar como un vector unitario de dimensión 1.

Ahora, en cambio, asociamos a cada nodo $i \in N$ un vector **unitario** $\boldsymbol{p}_i \in \mathbb{R}^n$, y consideramos la matrix $\boldsymbol{P}$ con estos vectores como filas.

Reemplazamos entonces la función objetivo $\boldsymbol{C} \bullet (\boldsymbol{x}\boldsymbol{x}^T)$ por $\boldsymbol{C} \bullet (\boldsymbol{P}\boldsymbol{P}^T)$ y las restricciones $x_i \in \{+1, -1\}$ por restricciones de los elementos de la diagonal de $\boldsymbol{P}\boldsymbol{P}^T$:

$$(\boldsymbol{P}\boldsymbol{P}^T)_{ii} = 1,$$

dado que un vector unitario $\boldsymbol{x}$ cumple $\boldsymbol{x} \cdot \boldsymbol{x} = 1$.

Relajación - Opción 2

Como $\boldsymbol{P}\boldsymbol{P}^T$ es semidefinida positiva, y cualquier matrix semidefinida positiva se puede factorizar de esta forma, vemos que mediante esta construcción obtenemos el mismo problema SDP de antes.

Considerando una matriz $\boldsymbol{P}$ donde todas las filas son de la forma $\boldsymbol{v}$ o $-\boldsymbol{v}$ para un vector unitario $\boldsymbol{v}$, vemos que este problema es efectivamente una relajación del problema (IQP).

Relación entre la relajación y el problema original

Como las formulaciones SDP son relajaciones del problema max-cut, el valor óptimo del problema SDP es una cota superior del valor óptimo del problema original.

Vamos a ver ahora que podemos usar la solución del problema SDP para obtener un corte razonablemente bueno.

Usamos la segunda relajación, y tomamos $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{P}^T$ una solución óptima de ese problema.

Si todas las filas de $\boldsymbol{P}$ fueran $+\boldsymbol{v}$ o $-\boldsymbol{v}$ para un vector $\boldsymbol{v}$, definimos el corte tomando en K todos los nodos para los cuales la fila correspondiente es $+\boldsymbol{v}$.

Ejemplo

Consideramos 5 puntos formando un pentágono, con costo 1 en los lados del pentágono y 0 en el resto de las aristas. Obtenemos la matriz $\boldsymbol{W} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$.

La solución óptima del problema SDP relajado es

$$\boldsymbol{X} = \begin{pmatrix} 1. & -0{,}81 & 0{,}31 & 0{,}31 & -0{,}81 \\ -0{,}81 & 1. & -0{,}81 & 0{,}31 & 0{,}31 \\ 0{,}31 & -0{,}81 & 1. & -0{,}81 & 0{,}31 \\ 0{,}31 & 0{,}31 & -0{,}81 & 1. & -0{,}81 \\ -0{,}81 & 0{,}31 & 0{,}31 & -0{,}81 & 1. \end{pmatrix},$$

que es una matriz de rango 2.

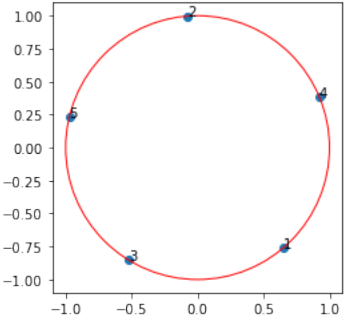
Ejemplo

Podemos calcular una descomposición $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{P}^T$, calculando la diagonalización $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^T$ y tomando $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{D}^{\frac{1}{2}}$. Obtenemos

$$\boldsymbol{P} = \begin{pmatrix} 0. & 0. & 0. & 0{,}65 & -0{,}76 \\ 0. & 0. & 0. & -0{,}08 & 1. \\ 0. & 0. & 0. & -0{,}52 & -0{,}85 \\ 0. & 0. & 0. & 0{,}92 & 0{,}38 \\ 0. & 0. & 0. & -0{,}97 & 0{,}23 \end{pmatrix}.$$

Ejemplo

Si graficamos las últimas dos coordenadas de los vectores filas, obtenemos 5 puntos en la circunferencia de radio 1.



Resulta razonable elegir un conjunto K tomando dos nodos vecinos en el gráfico.

Algoritmo Redondeo aleatorio de Goemans-Williamson

En la situación general, tomamos un hiperplano que divida a la esfera unitaria en dos mitades y tomamos en K a todos los nodos que quedan en una de estas dos mitades.

Más concretamente, para un vector aleatorio $\boldsymbol{g}$ de norma 1, definimos

$$K = \{i \in N : \boldsymbol{g} \cdot \boldsymbol{p}_i \geq 0\}.$$

Esto nos da un corte aleatorio, y podemos calcular la esperanza del valor del corte (veremos más adelante cómo tomar un vector aleatorio).

Relación entre la relajación y el problema original

La esperanza del valor del corte es

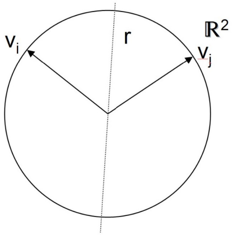
$$E[w(\delta(K))] = E\left[\sum_{i<j} w_{ij} \mathbb{1}_{(i,j)\in\delta(K)}\right] = \sum_{i<j} w_{ij} \Pr[(i,j) \in \delta(K)].$$

Observamos que las probabilidades no son independientes, pero no es necesaria independencia para distribuir la esperanza con respecto a la suma.

Relación entre la relajación y el problema original

Calculamos ahora $\Pr[(i,j) \in \delta(K)]$ para (i,j) fijo.

Consideramos el plano que pasa por el origen y contiene a los dos vectores $\boldsymbol{p}_i$ y $\boldsymbol{p}_j$.



La probabilidad de que un hiperplano separe a los dos vectores es la misma que la probabilidad de que un diámetro en este plano los separe.

Relación entre la relajación y el problema original

Esta probabilidad es

$$\Pr[(i,j) \in \delta(K)] = \frac{\angle(\boldsymbol{p}_i, \boldsymbol{p}_j)}{\pi} = \frac{\arccos(\boldsymbol{p}_i \cdot \boldsymbol{p}_j)}{\pi}.$$

Por lo tanto, el valor esperado del corte es

$$GW(K) = E[w(\delta(K))] = \sum_{i<j} w_{ij} \frac{\arccos(\boldsymbol{p}_i \cdot \boldsymbol{p}_j)}{\pi}.$$

Llamamos $SDP(K)$ al valor óptimo del problema SDP correspondiente a la solución $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{P}^T$.

Queremos comparar $E[w(\delta(K))]$ con este valor con $SDP(K)$.

Relación entre la relajación y el problema original

Recordemos que en la formulación original teníamos

$$\boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} = w(\delta(K)) = \sum_{i<j} w_{ij} \frac{1 - x_i x_j}{2}$$

y tomando $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$ obtuvimos

$$\boldsymbol{C} \bullet \boldsymbol{X} = w(\delta(K)) = \sum_{i<j} w_{ij} \frac{1 - x_{ij}}{2}$$

usando que $x_{ij} = x_i x_j$.

Por lo tanto, si tomamos $\boldsymbol{X} = \boldsymbol{P}^T \boldsymbol{P}$, las coordenadas de $\boldsymbol{X}$ son $x_{ij} = \boldsymbol{p}_i \cdot \boldsymbol{p}_j$ y por lo tanto

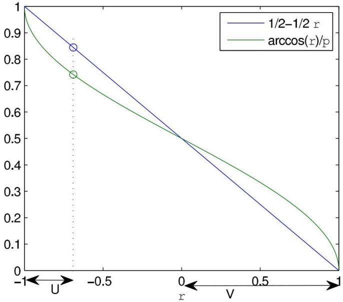
$$SDP(K) = \boldsymbol{C} \bullet (\boldsymbol{P}^T \boldsymbol{P}) = \sum_{i<j} w_{ij} \frac{1 - \boldsymbol{p}_i \cdot \boldsymbol{p}_j}{2}.$$

Relación entre la relajación y el problema original

Para comparar ambas expresiones término a término, consideramos

$$\rho = \boldsymbol{p}_i \cdot \boldsymbol{p}_j, \quad -1 \leq \rho \leq 1$$

y graficamos las funciones $f_1(\rho) = \frac{\arccos(\boldsymbol{p}_i \cdot \boldsymbol{p}_j)}{\pi}$ y $f_2(\rho) = \frac{1}{2} - \frac{1}{2}(\boldsymbol{p}_i \cdot \boldsymbol{p}_j)$.



Relación entre la relajación y el problema original

Calculando la mayor diferencia entre las dos funciones, obtenemos que

$$f_1(\rho) \geq 0{,}87854 f_2(\rho)$$

y por lo tanto:

$$GW(K) \geq 0{,}87854\ SDP(K),$$

y esto implica que existe algún corte $K^\star$ tal que $w(\delta(K^\star)) \geq 0{,}87854$ del valor óptimo del problema SDP.

Para el pentágono con todos los pesos de las aristas iguales a 1, la razón entre el valor óptimo del problema max-cut y la relajación SDP es aproximadamente $0{,}884$, por lo tanto la cota obtenida anteriormente esta cerca de la mejor cota que podemos obtener.