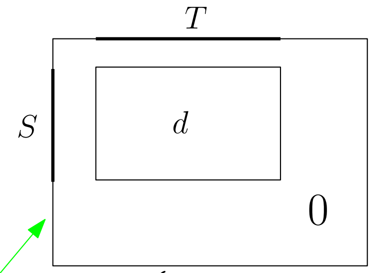
Quick approximation to matrices - *A. Frieze, R. Kannan*

predigested and performed by Marek Krčál



$R, |R| = m$ will denote set of rows, $C, |C| = n$ will denote set of columns, $R \times C$-real matrix A is $(A_{ij})_{(i,j)\in R\times C}, A_{ij} \in \mathbb{R}$

$$||A||_F := (\sum A_{ij}^2)^{1/2} \qquad ||A||_C := \max_{S\subseteq R T\subseteq C} |A_{S,T}| \text{ where } A_{S,T} := \sum_{(i,j)\in S\times T} A_{ij}$$

Cut decomposition of matrix A is given by R_j, C_j, d_j for $j = 1, \dots, s$ s.t.

$$\mathrm{CUT}(S,T,d)_{ij} := \begin{cases} d, & (i,j) \in S \times T \\ 0, & \text{otherwise} \end{cases}$$

$$A = D^{(1)} + D^{(2)} + \dots + D^{(s)} + W \text{ where } D^{(j)} = \mathrm{CUT}(R_j, C_j, d_j)$$

the decomposition's width (s), Error matrix (W), coefficient length is $(d_1^2 + \dots + d_s^2)^{1/2}$

Theorem 1. Let A be a real $R \times C$-matrix (R and S stands for set of rows and columns). There is $s < 0.56^2/\epsilon^2$: a cut decomposition of A can be constructed:
For every $t < 0.56^2/\epsilon^2$:

$$A = D^{(1)} + D^{(2)} + \dots + D^{(t)} + W^{(t)} \text{ where } D^{(j)} = \mathrm{CUT}(R_j, C_j, d_j)$$

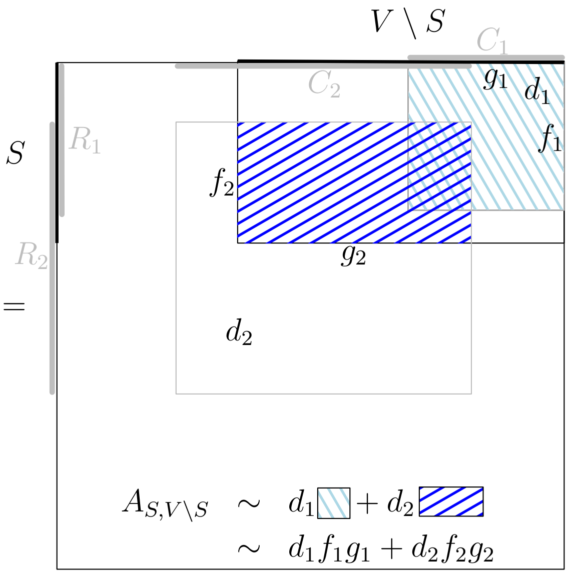
either
where $\forall S \subseteq R, \forall T \subseteq C : W^{(t)}_{S,T} \leq \epsilon\sqrt{|S||T|}||A||_F$ or $||W^{(t)}||_F^2 \leq (1 - 0.56^2\epsilon^2 t)||A||_F$

Constant time construction
For every $\delta > 0$ in time $2^{\tilde{O}(1/\epsilon^2)} \log\delta$ with probability $1-\delta$ the decomposition such as of width $s < 192/\epsilon^2$ and coefficient length $\sqrt{27}||A||_F/\sqrt{mn}$ can be constructed.

Theorem 2. Let $(A_{ij})_{i,j\in V}, A_{ij} \in [-1,1]$ be a matrix of edge weights of a complete graph. Then in time $2^{\tilde{O}(1/\epsilon^2)}\log(1/\delta)$ with probability $1-\delta$ we can find a cut $S^*, V \setminus S^*$ such that

$$A_{S^*,V\setminus S^*} \geq A_{S,V\setminus S} - \epsilon n^2 \text{ for every } S \subseteq V.$$

Proof.



- We use Theorem 1. to get a decomposition of A with error
$$||A - D^{(1)} - D^{(2)} - \dots - D^{(s)}||_C \leq \epsilon n ||A||_F/10 \leq \epsilon n^2/10$$
- $(D^{(1)} + \dots + D^{(s)})_{S,V\setminus S} = \sum d_j f_j g_j$ where $f_j = |S \cap R_j|$ and $g_j = |(V \setminus S) \cap C_j|$
- approximate: $\bar{f}_j := \lfloor f_j/\nu \rfloor \nu$ and $\bar{g}_j := \lfloor g_j/\nu \rfloor \nu$. We have
$$|\sum_{j=1}^{s} d_j f_j g_j - \sum_{j=1}^{s} d_j \bar{f}_j \bar{g}_j| \leq s\sqrt{27}(2n\nu - \nu^2) \leq \epsilon n^2/3$$
choose $\nu := \frac{\epsilon n}{9\sqrt{27}s}$ in order to get
- brute force: enumerate all $O(1/\epsilon^3)^{2s}$ possible values for $\bar{f}$ and $\bar{g}$, question whether for a given values of $\bar{f}, \bar{g}$ a cut S exists reduces to an integer program that we replace by its linear relaxation
  - Let $\mathcal{P}$ be the coarsest partition of V such that each R_j, C_j is a union of sets in $\mathcal{P}$
  - For every $\bar{f}, \bar{g}$ define the following IP:
    Find values x_P (represents $|S \cap P|$), $P \in \mathcal{P}$ subject to:
$$\begin{array}{rlll} 0 \leq & x_P & \leq |P| & P \in \mathcal{P} \\ \bar{f}_j \leq & \sum_{P\subseteq R_j} x_P & \leq \bar{f}_j + \nu & j = 1, \dots, s \\ \bar{g}_j \leq & \sum_{P\subseteq C_j}(|P| - x_P) & \leq \bar{g}_j + \nu & \end{array}$$
  - Find a feasible solution if it exists. Round down each value to the nearest integer below.