Lecture 20 10/31/2025

Swap Rounding for Spanning Trees

We previously saw pipage rounding as a way to convert a fractional solution $(x_1, x_2, \dots, x_n) \in [0,1]^n$ such that $\sum x_i = k$ into a **random** integer solution $(X_1, X_2, \dots, X_n)$

such that

(i) $X_i \in \{0,1\}$ and $E[X_i] = x_i$

(ii) $\sum_{i=1}^{n} X_i = k$

(iii) $X_1, X_2, \dots, X_n$ are negatively correlated.

Pipage rounding

1. While $\bar{x}$ is not fully integral do
   - Let $x_i, x_j \in (0,1)$ ie both fractional
   - Let $\varepsilon = \min\{x_i, x_j, 1-x_i, 1-x_j\}$
   - Toss a coin with prob $\frac{1}{2}$
   - If coin is heads

     $x_i \leftarrow x_i + \varepsilon$

     $x_j \leftarrow x_j - \varepsilon$
   - Else

     $x_i \leftarrow x_i - \varepsilon$

     $x_j \leftarrow x_j + \varepsilon$

   end while
2. Output $\bar{x}$.

Today the goal is to show a related but different scheme called swap rounding that works for fractional points in any matroid polytope. It is easier to see the combinatorics of this than the pipage rounding. We will explain this via spanning trees since they are familiar.

Defn: Let $G = (V, E)$ be a connected graph with m edges (can be a multigraph). Let $E = \{e_1, \dots, e_m\}$.

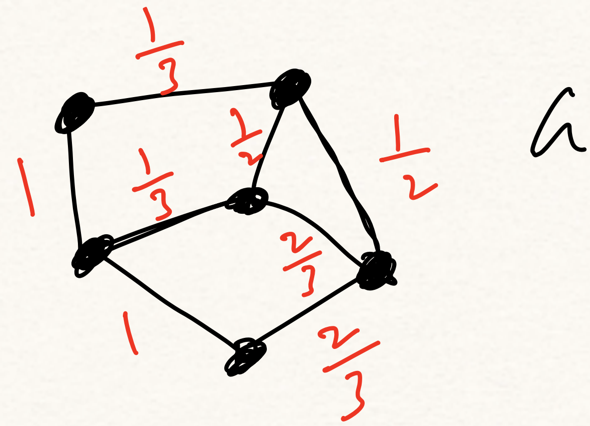
A vector $\bar{x} \in [0,1]^m$ is a <u>fractional spanning tree</u> of G if $\exists$ exist spanning trees $T_1, T_2, \ldots, T_h$ of G and coefficients $\lambda_1, \lambda_2, \ldots, \lambda_h \in [0,1]$ such that

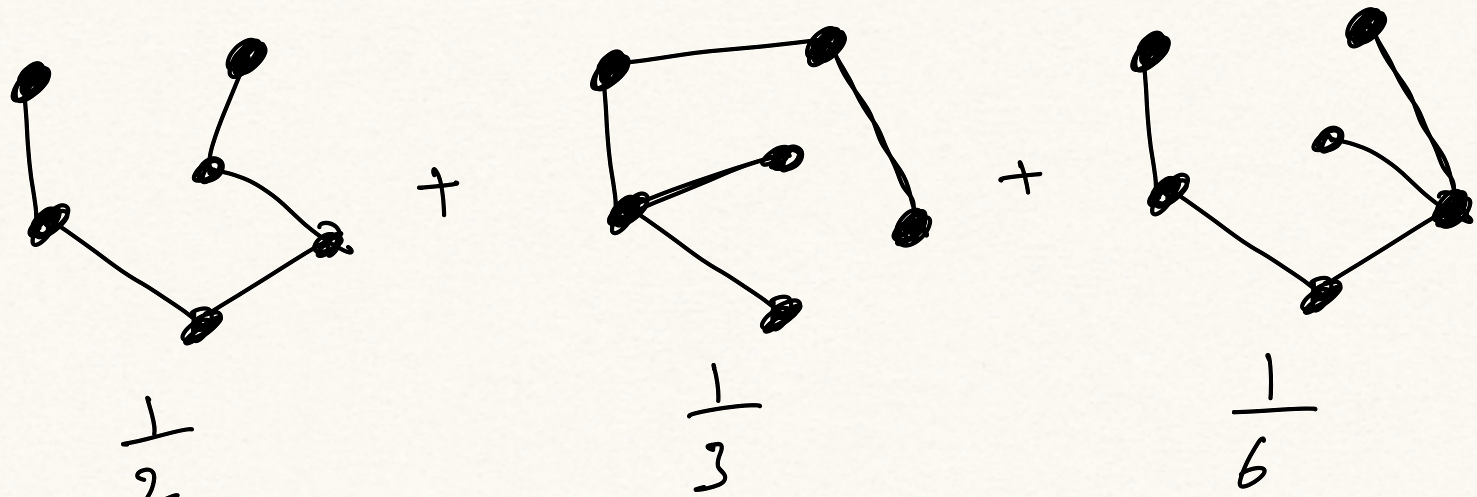
(i) $\sum_{i=1}^{h} \lambda_i = 1$ and

(ii) $\forall$ each edge e, $\bar{x}(e) = \sum_{i : e \in T_i} \lambda_i$

<u>Remark</u>: Note that $\bar{x}$ may correspond to different decompositions. Not necessarily unique.

Ex:





Suppose we are given a fractional spanning tree $\bar{x}$ along with a decomposition as $\sum_{i=1}^{k} \lambda_i 1_{T_i}$.

We want to round $\bar{x}$ into a random a spanning tree $(X_1, X_2, \dots, X_m)$.

A basic property we want is that $\forall e_i\ E[X_i = 1] = x_i$.

This is easy to accomplish. Simply pick a random tree from $T_1, \ldots, T_k$ where the probability that T_i is picked is equal to λ_i.

However $X_1, X_2, \ldots, X_m$ can be very correlated. We will now describe a way to do this in a different way. For this we will use a nice exchange property of spanning trees. This is more generally true of bases of any matroid.

Lemma: Let T_1 and T_2 be two spanning trees of a graph $G = (V, E)$. Suppose $E(T_1) \neq E(T_2)$. Then for *any* edge $e_1 \in E(T_1) - E(T_2)$ there is *an edge* $e_2 \in E(T_2) - E(T_1)$ such that $T_1 - e_1 + e_2$ *and* $T_2 - e_2 + e_1$ are both spanning trees.

Proof: exercise. $\square$

The idea in swap rounding is to use the exchange property to merge trees in a step by step way.

Merge $(\lambda_1, T_1, \lambda_2, T_2)$

- $\lambda_1 \in [0,1)$ $\lambda_2 \in [0,1)$
- While $E(T_1) \neq E(T_2)$ do
  - let $e_1 \in E(T_1) - E(T_2)$.
  - find $e_2 \in E(T_2) - E(T_1)$ such that $T_1 - e_1 + e_2$ and $T_2 - e_2 + e_1$ are spanning trees.
  - With prob $\frac{\lambda_1}{\lambda_1 + \lambda_2}$

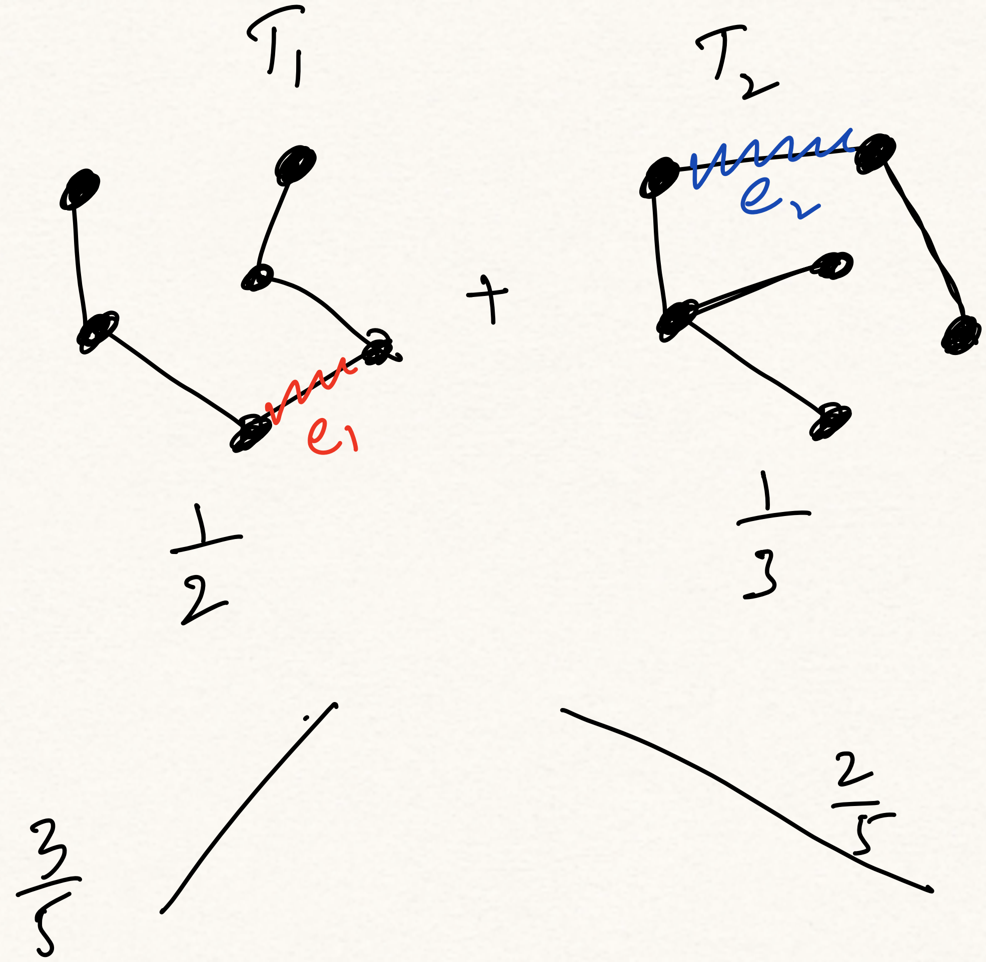
    $T_1 \leftarrow T_1 - e_1 + e_2$

    Else

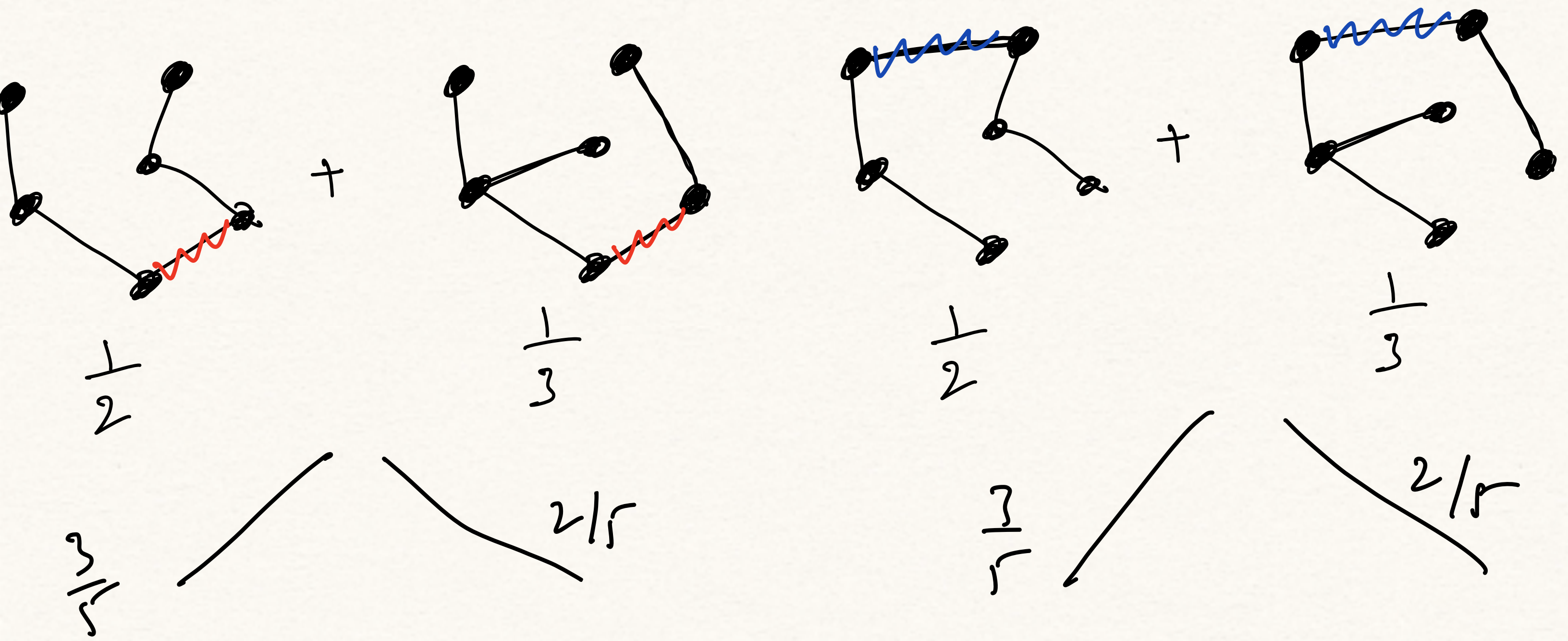
    $T_2 \leftarrow T_2 - e_2 + e_1$

- Output T_1. $(= T_2)$.

Ex:



T_1
T_2
e_1
e_2
+
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{3}{5}$
$\frac{2}{5}$



+
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{3}{5}$
$2/5$
+
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{3}{5}$
$2/5$

Merge finishes in $(n-1)$ steps.
Easy to implement with tree data structures. Swap rounding uses Merge to iteratively merge the trees one step at a time.

SwapRounding $(\lambda_1, \lambda_2, \ldots, \lambda_k, T_1, T_2, \ldots, T_k)$

- If $k=1$ return T_1
- Else
  $T = \text{Merge}(\lambda_1, T_1, \lambda_2, T_2)$
  return $((\lambda_1+\lambda_2), \lambda_3, \ldots, \lambda_k, T, T_3, \ldots T_k)$.

$$\bar{X} = \frac{1}{2} T_1 + \frac{1}{3} T_2 + \frac{1}{6} T_3$$

$$\frac{5}{6} T_{12} + \frac{1}{6} T_3$$

$$T_{123}.$$

It is easy to prove by induction that this form a martingale sequence. If $\bar{X}(t)$ is the vector after t steps then

(i) $E[\bar{X}(t+1)] = \bar{X}(t)$.

(ii) Only two coordinates of $\bar{X}$ change in each step.

Using above one can show that if the final vector is $(X_1, \ldots, X_m)$ then $X_1, \ldots, X_m$ are negatively correlated.

Lemma 4.1. *Let $\tau \in \mathbb{N}$ and let $\mathbf{X}_t = (X_{1,t}, \ldots, X_{n,t})$ for $t \in \{0, \ldots, \tau\}$ be a non-negative vector-valued random process with initial distribution given by $X_{i,0} = x_i$ with probability 1 $\forall i \in [n]$, and satisfying the following properties:*

1. $\mathbf{E}[\mathbf{X}_{t+1} \mid \mathbf{X}_t] = \mathbf{X}_t$ *for* $t \in \{0, \ldots, \tau\}$ *and* $i \in [n]$.
2. $\mathbf{X}_t$ *and* $\mathbf{X}_{t+1}$ *differ in at most two components for* $t \in \{0, \ldots, \tau - 1\}$.
3. *For* $t \in \{0, \ldots, \tau\}$, *if two components* $i, j \in [n]$ *change between* $\mathbf{X}_t$ *and* $\mathbf{X}_{t+1}$, *then their sum is preserved:* $X_{i,t+1} + X_{j,t+1} = X_{i,t} + X_{j,t}$.

Then for any $t \in \{0, \ldots, \tau\}$, *the components of* $\mathbf{X}_t$ *satisfy* $\mathbf{E}[\prod_{i \in S} X_{i,t}] \leq \prod_{i \in S} x_i \; \forall S \subseteq [n]$.

Proof. We are interested in the quantity $Y_t = \prod_{i \in S} X_{i,t}$. At the beginning of the process, we have $\mathbf{E}[Y_0] = \prod_{i \in S} x_i$. The main claim is that for each t, we have $\mathbf{E}[Y_{t+1} | \mathbf{X}_t] \leq Y_t$.

Let us condition on a particular configuration of variables at time t, $\mathbf{X}_t = (X_{1,t}, \ldots, X_{n,t})$. We consider three cases:

- If no variable $X_i, i \in S$, is modified in step t, we have $Y_{t+1} = \prod_{i \in S} X_{i,t+1} = \prod_{i \in S} X_{i,t} = Y_t$.
- If exactly one variable $X_i, i \in S$, is modified in step t, then by property 1 of the lemma:
$$\mathbf{E}[Y_{t+1} \mid \mathbf{X}_t] = \mathbf{E}[X_{i,t+1} \mid \mathbf{X}_t] \cdot \prod_{j \in S \setminus \{i\}} X_{j,t} = \prod_{j \in S} X_{j,t} = Y_t.$$
- If two variables $X_i, X_j, i, j \in S$, are modified in step t, we use the property that their sum is preserved: $X_{i,t+1} + X_{j,t+1} = X_{i,t} + X_{j,t}$. This also implies that
$$\mathbf{E}[(X_{i,t+1} + X_{j,t+1})^2 \mid \mathbf{X}_t] = (X_{i,t} + X_{j,t})^2. \tag{1}$$
On the other hand, the value of each variable is preserved in expectation. Applying this to their difference, we get $\mathbf{E}[X_{i,t+1} - X_{j,t+1} \mid \mathbf{X}_t] = X_{i,t} - X_{j,t}$. Since $\mathbf{E}[Z^2] \geq (\mathbf{E}[Z])^2$ holds for any random variable, we get
$$\mathbf{E}[(X_{i,t+1} - X_{j,t+1})^2 \mid \mathbf{X}_t] \geq (X_{i,t} - X_{j,t})^2. \tag{2}$$
Combining (1) and (2), and using the formula $XY = \frac{1}{4}((X+Y)^2 - (X-Y)^2)$, we get
$$\mathbf{E}[X_{i,t+1} X_{j,t+1} \mid \mathbf{X}_t] \leq X_{i,t} X_{j,t}.$$
Therefore,
$$\mathbf{E}[Y_{t+1} \mid \mathbf{X}_t] = \mathbf{E}[X_{i,t+1} X_{j,t+1} \mid \mathbf{X}_t] \cdot \prod_{k \in S \setminus \{i,j\}} X_{k,t} \leq \prod_{k \in S} X_{k,t} = Y_t,$$

as claimed. By taking expectation over all configurations $\mathbf{X}_t$ we obtain $\mathbf{E}[Y_{t+1}] \leq \mathbf{E}[Y_t]$. Consequently, $\mathbf{E}[\prod_{i \in S} X_{i,t}] = \mathbf{E}[Y_t] \leq \mathbf{E}[Y_{t-1}] \leq \ldots \leq \mathbf{E}[Y_0] = \prod_{i \in S} x_i$, as claimed by the lemma. ∎

Continuous Extensions of Set Functions

Suppose we have a set function

$f : 2^N \to R$ that is real-valued

for each subset $S \subseteq N$ $f(S)$
is the value of S.

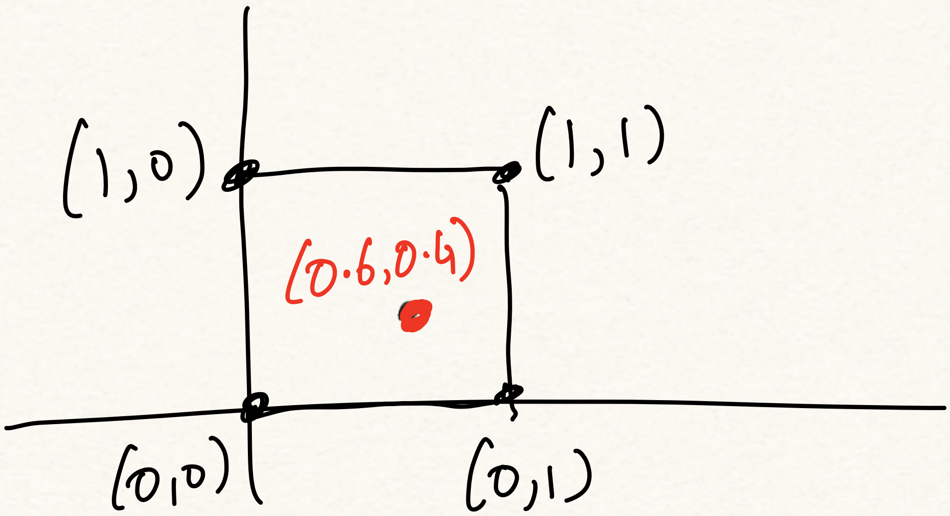
Many time it is useful to extend
f to fractional values.

$g(x_1, x_2, \ldots, x_n)$ where $x_i \in [0,1]$
$\forall i \in [n]$

How should we define it?
We know $f(x_1, \ldots, x_n)$ if $\bar{x} \in \{0,1\}^n$

We can use probability distribution view for interpolation.

given $\bar{x} \in [0,1]^n$ there are multiple ways of expressing $\bar{x}$ as a convex combination of the corners



Let α_S $S \subseteq N$ s.t. that

$$\sum_{S \ni i} \alpha_S = x_i \quad \forall i \in N$$

$$\sum \alpha_S = 1$$

$$\alpha_S \geq 0 \quad \forall S$$

This is the space of all convex combinations of sets that can express $\bar{x}$. We can think of α_S as a probability distribution.

Convex closure

We obtain one extension f^- called the convex closure by choosing the distribution that minimizes the expected value of f.

$$f^-(\bar{x}) = \min \sum_{S \subseteq N} \alpha_S f(S)$$

$$\sum_{S \subseteq N} \alpha_S = 1$$

$$\sum_{S \ni i} \alpha_S = x_i \quad i \in N$$

$$\alpha_S \geq 0 \quad S \subseteq N.$$

Claim: $f^-(x)$ is a convex function for any set function f.

Proof: Exercise.

Similarly we can define concave closure f^+ if we change min to max in the above LP.

Two other extensions are the Multilinear extension and the Lovasz extension.