Cleanup notes on max-cut.

. **Max-Cut**:

Given $G = (V, E)$, $w : E \to Z^+$,

Determine $S, \bar{S}$ s.t. $S \cup \bar{S} = V$

$S \cap \bar{S} = \phi$

$w(S, \bar{S})$ is **maximised**

. Randomized strategy.

- randomly assign $-1, 1$ to verts
- $S = \{$ Verts w/ 1 $\}$
- $\bar{S} = \{$ .. .. $-1\}$

In this case, $P[u\text{-}v \text{ crosses cut}]$

$$= P[u=1, v=-1] + P[u=-1, v=1]$$

$$= 1/2.$$

now, the value of the cut is

$$Z_{y_r} = \frac{1}{2} \sum_{i>j} w_{ij} [1 - y_i y_j]$$

$$E[Z_{yr}] = \frac{1}{2}\sum_{i>j} w_{ij}\,(1 - E[y_i y_j])$$

$$E[y_i y_j] = P[y_i y_j = 1]\cdot 1 + P[y_i y_j = -1]\cdot -1$$

$$= 0$$

$$\therefore E[Z_{yr}] = \frac{1}{2}\sum_{i>j} w_{ij}$$

now the value of prob is value of best cut write Z_{opt}.

$$Z_{opt} \leq \sum_{i>j} w_{ij}$$ (because weights are non-neg)

$$\therefore \boxed{E[Z_{yr}] = \frac{1}{2}\sum_{i>j} w_{ij} \geq \frac{1}{2} Z_{opt}}$$

G+W bound:

we solve

$$\max \ \frac{1}{2} \sum_{i>j} w_{ij} (1 - v_i \cdot v_j)$$

$$\text{s.t.} \quad v_i \cdot v_j = 1$$

(which is an SDP). value of this is Z_P^*

- we now choose a random vector $\underline{r}$
- at vert i, if $v_i \cdot r > 0$, $i \to S$
  $i \to \bar{S}$

Now:

$$P_r[\, i,j \in \text{cut} \,] = \frac{\cos^{-1}(v_i \cdot v_j)}{\pi} \qquad \text{(elementary geometry)}$$

write W for the value of the cut.

$$E[W] = \frac{1}{\pi} \sum_{i>j} w_{ij} \cdot \cos^{-1}(v_i \cdot v_j)$$

Now, for $\alpha \approx 0.878\ldots$

$$\frac{\theta}{\pi} \geq \alpha \cdot \frac{1}{2}(1-\cos\theta)$$

[which we can get by minimizing $\frac{\theta}{\pi} \cdot \frac{2}{(1-\cos\theta)}$]

<u>but</u> write $\theta_{ij} = \cos^{-1}(v_i \cdot v_j)$

we have

$$E[W] = \frac{1}{\pi} \cdot \sum_{i>j} w_{ij} \cos^{-1}(v_i \cdot v_j)$$

$$= \sum_{i>j} w_{ij} \frac{\theta_{ij}}{\pi}$$

$$\geq \alpha \cdot \sum_{i>j} w_{ij} (1-\cos\theta_{ij}) = \alpha Z_p^*$$

$$\alpha Z_p^* \geq \alpha Z_{opt}$$

So $E[W] \geq \alpha Z_{opt}$

Doubly stochastic matrices and balancing

- Recall M is doubly stochastic if
  - $M_{ij} \geq 0$
  - $\sum_i M_{ij} = \sum_j M_{ij} = 1$

(Notice this means that M is square.)

- Recall.
  - Convex hull of permutation matrices = doubly stochastic matrices

    (easy).
  - Vertices of doubly stochastic matrix polytope = permutation matrices

    (Deep, non-obvious; Birkhoff).
- Above establishes how d.s. matrices could be interesting
  - also, matching paper.

Procedure to produce d.s. matrix from M.

Iterate

$$M^{*}_{ij} = \frac{M^{(n)}_{ij}}{\sum_j M^{(n)}_{ij}}$$

$$\hat{M}_{ij} = \frac{M^{*}_{ij}}{\sum_i M^{*}_{ij}}$$

~~and let~~

$$M^{(n+1)}_{ij} = \hat{M}_{ij}$$

and keep going, till convergence.

Facts:

- if procedure converges, it converges to unique d.s. matrix.
- Zeros may cause a failure to converge (conditions under which it converges are known, see papers)
- it doesn't help to make zero's into ε (eg next page)

<u>example</u>:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{approx as} \quad \begin{pmatrix} \varepsilon & \varepsilon & 1 \\ \varepsilon & \varepsilon & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

converges to

$$\begin{pmatrix} 0.25 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.5 \\ .5 & .5 & 0 \end{pmatrix} \quad \varepsilon \to 0$$

<u>but</u>

$$\cancel{0} \quad \cancel{0} \quad 1$$
$$\cancel{0}$$

$$\begin{matrix} \varepsilon & \varepsilon^2 & 1 \\ \varepsilon & \varepsilon^2 & 1 \\ 1 & 1 & 1 \end{matrix} \quad \text{conv to}$$

$$\begin{pmatrix} .5 & 0 & .5 \\ .5 & 0 & .5 \\ 0 & 1 & 0 \end{pmatrix} \quad \varepsilon \to 0 \quad !$$

This procedure has some quite surprising applications.

Sinkhorn can solve Sudoku!

Sudoku : $N \times N$ grid of cells

- N blocks of N elements
- into each cell, insert $1 \cdots N$

s.t. · exactly 1 in each
- row
- col
- block

- N is usually 9.
- There is initial data (and the solver must fill in the rest).
- There are $3N$ vector constraints
- have a doubly-stochastic flavor
- e.g.

write $$S_{ij}^{l} = \begin{cases} 1 & \text{if } ij \text{ the cell has } l \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\sum_i S_{ij}^{\ell} = 1$$

$$\sum_j S_{ij}^{\ell} = 1$$

$$\sum_{i,j \in \text{Block}} S_{ij}^{\ell} = 1 \quad \text{for each block.}$$

$$\sum_\ell S_{ij}^{\ell} = 1 \quad \text{(only one \# per table location.}$$

<u>Procedure</u>:

~~row~~ ~~col~~ ~~block~~ cell

- write $p^m = [\, p(\cancel{\ell}_m = 1),\ p(loc = 2), \ldots p(loc = N)]$

(which we interpret as probabilities, but...

- Now imagine a col / row / block r matrix

values →

$$\begin{array}{c} \\ \text{cells} \downarrow \end{array} \begin{bmatrix} p^{m_1} & \\ p^{m_2} & \\ \cdot & \\ & \end{bmatrix} = Q_\gamma$$

at the right answer, all Q^r would be permutations (6)

- for prescribed data, we can adjust the p-vector (i.e. it must be $(0 \cdots 1 \cdots 0)$
  known value.
- This means that Q^r at least are doubly stochastic.
- idea:
  - construct each row Q^r
    - straighten
  - " " col
    - straighten
  - " " block
    - straighten
  - stop on convergence.

Each balancing step does not move farther from soln in appropriate distance.

Notation:

$\hat{S}_{ij}^{l}$ for a solution

$$\hat{S}_{ij}^{l} = \begin{cases} 1 & i,j\text{'th cell} = l \\ 0 & \text{otherwise} \end{cases}$$

Notice $\hat{S}_{ij}^{l}$ has balance properties

Write $\hat{g}_{ij}^{l}$ for the n'th est of $\hat{S}_{ij}^{l}$

~~Notice~~

Start with $\overset{\circ}{g}_{ij}^{l}$ st $\sum_{l} \overset{\circ}{g}_{ij}^{l} = 1$

Each step of each stage does not increase $D(\hat{S} \| g) = \sum_{ij\ell} \hat{S}_{ij}^{\ell} \log \frac{\hat{S}_{ij}^{\ell}}{g_{ij}^{\ell}}$

proof: (case-by-case)

eg: col balancing for a row matrix

. matrix is Q

~~$Q_{j\ell} = \hat{S}_{ij}^{\ell}$~~

. for k'th row, $Q_{j\ell} = \hat{S}_{kj}^{\ell}$ (k fixed)

and col balancing takes

$$Q_{j\ell} \rightarrow \frac{Q_{j\ell}}{\sum_j Q_{j\ell}} = \frac{Q_{j\ell}}{X_\ell}$$

i.e. this gets

$$\overset{(n+1)}{g}{}^{l}_{ij} = \overset{(n)}{g}{}^{l}_{ij} \qquad \text{for } i \neq k$$

$$\overset{(n+1)}{g}{}^{l}_{kj} = \frac{\overset{(n)}{g}{}^{l}_{kj}}{X_l} \qquad i = k$$

now

$$D(\hat{s} \,|\, \overset{(n+1)}{g}) \leq D(\hat{s} \,|\, \overset{(n)}{g})$$

$$= \sum_{ijl} \hat{s}^{l}_{ij} \log \frac{\hat{s}^{l}_{ij}}{\overset{(n)}{g}{}^{l}_{ij}} + \sum_{jl} \hat{s}^{l}_{kj} \log X_l$$

the change is

$$\sum_{jl} \hat{s}^{l}_{kj} \log X_l \leq \sum_{jl} \hat{s}^{l}_{kj} (X_l - 1)$$

because $(\log x \leq x - 1)$

$$= \sum_{jl} \hat{s}^{l}_{kj} X_l - N$$

$$= N - N$$

$$= 0$$

Other cases follow same line □

In practice, works OK.