1 Uniform Sampling for Spanning Trees

Consider an unweighted undirected graph.

A corollary of a core argument from the uniform sampling for matrix approximation result (Lecture 22 on Nov 5) is that if we sample half the edges, the expected number of edges with endpoints in different connected components is $O(n)$.

Give a combinatorial proof of this fact. No linear algebra allowed :-).

2 Leverage Scores and $\ell_1$ Lewis Weights

Lewis weights are sampling probabilities that preserve $\|Ax\|_p$ (described in Lecture 23 on Nov 10). Specifically, the $\ell_1$-Lewis weights for a matrix $A$ with rows $a_i$ were defined as:

$$w_i = \sqrt{a_i (A^T W^{-1} A)^{-1} a_i^T}.$$

Also, recall that the statistical leverage scores are:

$$\tau_i = a_i (A^T A)^{-1} a_i^T.$$

Show that $O(n^{1/2}) \tau_i \geq w_i$. This implies that sampling $O(n^{1/2} d \log n)$ rows by statistical leverage scores still gives $\tilde{A}$ such that $\|Ax\|_1 \approx \|\tilde{A}x\|_1 \forall x$.

3 Taylor Series Approximation of $\ln x$

Show that if $0.9 \leq x \leq 1.1$, we have

$$x - 1 - 10x^2 \leq \ln x \leq x - 1 + 10x^2.$$
4 Long Steps

Consider minimizing the log-barrier function:

$$\min_{Ax=b} c^T x - 1^T \log x,$$

with gradient

$$\nabla = c - \frac{1}{x}.$$

Interior point algorithms rely on the fact that if we have $\Delta$ such that $\Delta^T \nabla > 0.2$ and $\|X^{-1} \Delta\|_2^2 < 0.1$, we can reduce the objective by 0.1.

We will now consider taking bigger steps: show that if we have $\Delta$ such that

- $\Delta^T \nabla > k$, and
- $\|X^{-1} \Delta\|_2^2 < 0.1$, and
- $x_i^{-1} \Delta_i \leq 2/k$ for all $i$,

then we can reduce the objective by $\Theta(k^2)$.

5 ‘Slower’ Variant of Interior Point Method

Consider the following variant of the path following algorithm described in Lecture 25 on Nov 17. Instead of multiplying $t$ by $1 + 1/\sqrt{m}$ when the descent step says we’re close to the optimum, double $t$ instead.

Show that $O(m)$ steps suffices between each doubling of $t$. Specifically, show that if $x$ is close to optimum for $\Phi(t)$, then

$$\Phi(2t)(x) \leq \min_{x'} \Phi(2t)(x') + O(m),$$

where $\Phi(t)(x) = tc^T x - 1^T \log x$.

This analysis should be easier than the one done in class, as it more closely matches Karmarkar’s original algorithm.

6 Uniformly Sampling from the $n$-Dimensional Simplex

Show that the following routine generates a point uniformly at random from the $n$-dimensional simplex:

$$K = \left\{ x : x \geq 0, \sum_i x_i = 1 \right\}.$$
1. Generate $y_1 \ldots y_{n-1}$ uniformly random from $[0, 1]$, aka. $y \sim U[0, 1]^{n-1}$.

2. Sort them, and let $y_0 = 0$, $y_n = 1$.

3. Set $x_i = y_i - y_{i-1}$.

7 Section of a Ball

Let $B_n$ be the unit $n$ dimensional ball, aka.

\[ B_n = \{ x : \|x\|_2 \leq 1 \} \, . \]

Show that there exists absolute constants $c_1 > 0$ and $c_2 < 1$ such that:

\[ \Pr_{x \in \text{Uniform}(B_n)} \left[ x_1 \geq \frac{c_1}{\sqrt{n}} \right] \leq c_2. \]

Where Uniform($B_n$) is the uniform distribution over $B_n$. 
