1 New Random Edge Weights at Each Iteration

Define the PageRank iteration $x^{(k+1)} = \alpha(P + Q_k)x^{(k)} + (1 - \alpha)v$, where each Q_k is independently drawn from the probability distribution function f, and the other variables are defined as usual. We require that f be defined such that $P + Q_k$ is always a valid PageRank matrix. For example, P and $P + Q_k$ must be column stochastic, so the elements of any column of Q_k must sum to zero. Furthermore, without loss of generality, assume that $\mathrm{E}[f] = 0$. (If $\mathrm{E}[f]$ was not zero, we could redefine P to be $P + \mathrm{E}[f]$ and f to be $f - \mathrm{E}[f]$.)

By the linearity of expectation, $\mathrm{E}\left[x^{(k+1)}\right] = \alpha P \mathrm{E}\left[x^{(k)}\right] + (1 - \alpha)v$, and of course $\mathrm{E}\left[x^{(0)}\right] = x^{(0)}$, so computing $\mathrm{E}\left[x^{(k+1)}\right]$ is equivalent to a traditional PageRank problem without randomness. Determining the likely accuracy of this iteration is more complicated. (Note: To ease notation, I use $y = x^{(k+1)}$ and $x = x^{(k)}$ for the rest of this section.) My first attempt was to compute $\mathrm{Var}(\|y\|_2)$, as follows:

$$\begin{aligned}
\mathrm{Var}(\|y\|) &= \mathrm{E}\left[\|y\|^2\right] - \mathrm{E}[\|y\|]^2 \\
&= \mathrm{E}\left[\sum_{i=1}^{n} y_i^2\right] - \mathrm{E}[\|y\|]^2 \\
&= \left(\sum_{i=1}^{n} \mathrm{E}\left[y_i^2\right]\right) - \mathrm{E}[\|y\|]^2
\end{aligned}$$

I don't know how to compute the last term, but knowing $\mathrm{E}\left[y_i^2\right]$ would allow us to compute $\mathrm{Var}(y_i)$, since $\mathrm{Var}(y_i) = \mathrm{E}\left[y_i^2\right] - \mathrm{E}[y_i]^2$, and we know how to compute $\mathrm{E}[y_i]$. More generally, I derived an algorithm for computing $\mathrm{E}\left[yy^T\right]$ given $\mathrm{E}\left[xx^T\right]$. (This allows us to recursively compute $\mathrm{E}\left[yy^T\right]$.) We can read the expected values of the squares of the elements of y off the diagonal of this matrix. Furthermore, $\mathrm{E}\left[yy^T\right] - \mathrm{E}[y]\,\mathrm{E}[y]^T$ is the covariance matrix of the elements of y, since $\left(\mathrm{E}\left[yy^T\right] - \mathrm{E}[y]\,\mathrm{E}[y]^T\right)_{i,j} = \mathrm{E}[y_i y_j] - \mathrm{E}[y_i]\,\mathrm{E}[y_j] = \mathrm{Cov}(y_i, y_j)$.

Suppose we're given Q, and let $A = P + Q$, so $y = \alpha A x + (1 - \alpha)v$. Then,

$$\begin{aligned}
\mathrm{E}\left[yy^T \mid Q\right] &= \mathrm{E}\left[(\alpha A x + (1 - \alpha)v)\left(\alpha x^T A^T + (1 - \alpha)v^T\right) \mid Q\right] \\
&= \alpha^2 A \mathrm{E}\left[xx^T\right] A^T + \alpha(1 - \alpha)(A \mathrm{E}[x]\, v^T + v \mathrm{E}[x]^T A^T) + (1 - \alpha)^2 vv^T
\end{aligned}$$

Since $\mathrm{E}[Q] = 0$ and $\mathrm{E}[A] = \mathrm{E}[P] + \mathrm{E}[Q] = P$, we have the following:

$$\mathrm{E}\left[yy^T\right] = \alpha^2 P \mathrm{E}\left[xx^T\right] P^T + \alpha^2 \mathrm{E}\left[Q \mathrm{E}\left[xx^T\right] Q^T\right] + \alpha(1 - \alpha)(P \mathrm{E}[x]\, v^T + v \mathrm{E}[x]^T P^T) + (1 - \alpha)^2 vv^T$$

Since we know $\mathrm{E}[x]$ and $\mathrm{E}\left[xx^T\right]$, the only term we don't know is $\mathrm{E}\left[Q \mathrm{E}\left[xx^T\right] Q^T\right]$. For all integers q, r, s, t from 1 to n, let $c_{q,r,s,t} = \mathrm{Cov}(f_{q,r}, f_{s,t}) = \mathrm{E}[f_{q,r} f_{s,t}] - \mathrm{E}[f_{q,r}]\,\mathrm{E}[f_{s,t}] = \mathrm{E}[f_{q,r} f_{s,t}]$, the covariance of two elements of the output of f.

(These values can be precomputed before any iterations of this algorithm.) Then $(Q\mathrm{E}\left[xx^T\right]Q^T)_{q,s} = Q_{q,:}\mathrm{E}\left[xx^T\right](Q_{s,:})^T = \sum_{r,t=1}^{n} Q_{q,r}\mathrm{E}\left[xx_T\right]_{r,t} Q_{s,t}$, and its expected value is $\sum_{r,t=1}^{n} c_{q,r,s,t}\mathrm{E}\left[xx^T\right]_{r,t}$. The entire matrix $\mathrm{E}\left[Q\mathrm{E}\left[xx^T\right]Q^T\right]$ can be computed by iterating over r and t from 1 to n.

Some remaining problems include finding a closed form for this algorithm, and/or an approximation algorithm that makes it easier to understand how the variance of f results in variance in the PageRank vector, especially as k becomes large. I decided to hold off on that for now in case what I've done so far is on the wrong track.

2 Approximations Regarding Random Edge Weights in a Direct Solution

Suppose we have a PageRank Problem $Mx = b$, but now we add randomness, so we try to solve $(M - E)x_e = b$, where E is random and has expected value 0. Assume that E is small enough so that $(M - E)^{-1} = \sum_{k=0}^{\infty} M^{-1}(EM^{-1})^k$. If E is guaranteed to be sufficiently small, then $(M - E)^{-1} \approx M^{-1} + M^{-1}EM^{-1}$. ($M^{-1}EM^{-1}EM^{-1}$ is a good estimate of the error of this first order approximation, for sufficiently small E.) This allows us to approximate the effect of randomness on the solution:

$$
\begin{aligned}
x_e - x &= (M - E)^{-1}b - x \\
&= (M - E)^{-1}Mx - x \\
&= \Big((M - E)^{-1}M - I\Big)\, x \\
&\approx \big((M^{-1} + M^{-1}EM^{-1})M - I\big)\, x \\
&= (I + M^{-1}E - I)x \\
&= M^{-1}Ex
\end{aligned}
$$

It follows that $\mathrm{E}\left[x_e - x\right] \approx M^{-1}\mathrm{E}\left[E\right]x = 0$, and also, $\mathrm{E}\left[(M - E)^{-1}\right] \approx M^{-1} + M^{-1}\mathrm{E}\left[E\right]M^{-1} = M^{-1}$. Regardless of the approximation used, knowing $\mathrm{E}\left[(M - E)^{-1}\right]$ allows us to find the expected solution of the PageRank problem, since $\mathrm{E}\left[(M - E)^{-1}b\right] = \mathrm{E}\left[(M - E)^{-1}\right]b$. If instead we want a second order approximation of $(M - E)^{-1}$, we can compute its expected value as follows:

$$
\begin{aligned}
\mathrm{E}\left[(M - E)^{-1}\right] &\approx \mathrm{E}\left[M^{-1} + M^{-1}EM^{-1} + M^{-1}EM^{-1}EM^{-1}\right] \\
&= M^{-1} + M^{-1}\mathrm{E}\left[EM^{-1}E\right]M^{-1}
\end{aligned}
$$

$\mathrm{E}\left[EM^{-1}E\right]$ depends on the covariances of the elements of E, and can be computed in a manner similar to that in the last paragraph of the previous section. Similarly, a third-order approximation also depends on $M^{-1}\mathrm{E}\left[EM^{-1}EM^{-1}E\right]M^{-1}$, which depends on the expected values of products of three elements of E. The same pattern holds for higher-order approximations.

The variance of the first-order approximation of $(M-E)^{-1}$ is $\mathrm{Var}\left(M^{-1}+M^{-1}EM^{-1}\right) = \mathrm{Var}\left(M^{-1}EM^{-1}\right) = \mathrm{E}\left[(M^{-1}EM^{-1})^2\right] - \mathrm{E}\left[M^{-1}EM^{-1}\right]^2 = \mathrm{E}\left[M^{-1}EM^{-2}EM^{-1}\right] - 0^2 = M^{-1}\mathrm{E}\left[EM^{-2}E\right]M^{-1}$, which depends on the covariances of the elements of E. The variance of the second-order approximation can be computed as follows:

$$
\begin{aligned}
\mathrm{Var}\left(\cdot\right) &= \mathrm{Var}\left(M^{-1}+M^{-1}EM^{-1}+M^{-1}EM^{-1}EM^{-1}\right)\\
&= \mathrm{Var}\left(M^{-1}EM^{-1}+M^{-1}EM^{-1}EM^{-1}\right)\\
&= M^{-2}\mathrm{Var}\left(E+EM^{-1}E\right)M^{-2}\\
&= M^{-2}\left(\mathrm{E}\left[(E+EM^{-1}E)^2\right] - \mathrm{E}\left[E+EM^{-1}E\right]^2\right)M^{-2}\\
&= M^{-2}\left(\mathrm{E}\left[E^2\right]+\mathrm{E}\left[E^2M^{-1}E\right]+\mathrm{E}\left[EM^{-1}E^2\right]+\mathrm{E}\left[EM^{-1}E^2M^{-1}E\right]-\mathrm{E}\left[EM^{-1}E\right]^2\right)M^{-2}
\end{aligned}
$$

If the kth order approximation of $(M-E)^{-1}$ is $\sum_{i=0}^{k}(M^{-1}E)^iM^{-1}$, then its error is $\sum_{i=k+1}^{\infty}(M^{-1}E)^iM^{-1} = (M^{-1}E)^{k+1}\sum_{i=0}^{\infty}(M^{-1}E)^iM^{-1} = (M^{-1}E)^{k+1}(M-E)^{-1}$. We can bound the relative error of this approximation as follows:

$$
\begin{aligned}
\frac{\|(M^{-1}E)^{k+1}(M-E)^{-1}\|}{\|(M-E)^{-1}\|} &\leq \frac{(\|M^{-1}\|\|E\|)^{k+1}\|(M-E)^{-1}\|}{\|(M-E)^{-1}\|}\\
&= (\|M^{-1}\|\|E\|)^{k+1}
\end{aligned}
$$

If $M = I-\alpha P$, then $\|M^{-1}\| = \|\sum_{k=0}^{\infty}(\alpha P)^k\| \leq \sum_{k=0}^{\infty}(\alpha\|P\|)^k = \frac{1}{1-\alpha\|P\|}$, so $(\|M^{-1}\|\|E\|)^{k+1} \leq \left(\frac{\|E\|}{1-\alpha\|P\|}\right)^{k+1}$. Since the P is column stochastic in the PageRank problem, it's 1-norm is 1, so the relative error in the 1-norm is bounded by $\left(\frac{\|E\|_1}{1-\alpha}\right)^{k+1}$.

3 Miscellaneous Observations

Most of these observations don't seem useful at the moment, but it seemed like a good idea to keep them in case they prove useful later.

The Sherman Morrison Woodbury formula gives $(M+E)^{-1} = M^{-1} - M^{-1}(E^{-1}+M^{-1})^{-1}M^{-1}$, which looks similar to the first order approximation for $(M+E)^{-1}$.

Suppose we have a PageRank iteration $x_{k+1} = \alpha Px_k + (1-\alpha)v$. Let $e_k = x_k - x_*$, where x_* is the fixed point of the iteration. Subtracting the equation $x_* = \alpha Px_* + (1-\alpha)v$ from the original iteration equation yields $e_{k+1} = \alpha Pe_k$, so $e_k = (\alpha P)^k e_0$.

Suppose we have a PageRank iteration $x_{k+1} = \alpha P x_k + (1-\alpha)v$, with initial guess x_0. I prove by induction that $x_k = (\alpha P)^k x_0 + (I-(\alpha P)^{k+1})(I-\alpha P)^{-1}(1-\alpha)v$. A simple computation shows this is true for $k=0$. If it's true for k, then

$$\begin{aligned}
x_{k+1} &= \alpha P x_k + (1-\alpha)v \\
&= \alpha P \left((\alpha P)^k x_0 + (I-(\alpha P)^{k+1})(I-\alpha P)^{-1}(1-\alpha)v\right) + (1-\alpha)v \\
&= (\alpha P)^{k+1} x_0 + \left(\alpha P (I-\alpha P)^{-1} - (\alpha P)^{k+1}(I-\alpha P)^{-1} + I\right)(1-\alpha)v \\
&= (\alpha P)^{k+1} x_0 + (\alpha P - (\alpha P)^{k+1} + I - \alpha P)(I-\alpha P)^{-1}(1-\alpha)v \\
&= (\alpha P)^{k+1} x_0 + (I-(\alpha P)^{k+2})(I-\alpha P)^{-1}(1-\alpha)v
\end{aligned}$$

Suppose we have $x = \alpha P x + (1-\alpha)v$. Note that $(I-\alpha P)^{-1} = \sum_{k=0}^{\infty} (\alpha P)^k = I + \sum_{k=1}^{\infty} \alpha^k P^k$. Therefore $x = (I-\alpha P)^{-1}(1-\alpha)v = (I + \sum_{k=1}^{\infty} \alpha^k P^k)(1-\alpha)v = v + \sum_{k=1}^{\infty} \alpha^k (P^k v - P^{k-1} v) = v + \sum_{k=1}^{\infty} \alpha^k P^{k-1}(P-I)v = \left(I + \sum_{k=1}^{\infty} (\alpha^k P^{k-1})(P-I)\right) v = (I + \alpha(I-\alpha P)^{-1})(P-I)v$.

$(M-E)^{-1} = M^{-1} + M^{-1}E(M-E)^{-1}$, so $\mathrm{E}\left[(M-E)^{-1}\right] = M^{-1} + M^{-1}\mathrm{E}\left[E(M-E)^{-1}\right]$. This would be useful if E is drawn from a probability distribution for which $\mathrm{E}\left[E(M-E)^{-1}\right]$ is easier to compute than $\mathrm{E}\left[(M-E)^{-1}\right]$.