A NONLINEAR PROBLEM IN LINEAR ALGEBRA AUGUST

RAYMOND FRIEND

Abstract. Undergraduate Research is something to be revered, I swear.

Contents

1. Introduction

2. Problem, Problems, Problemsen, Problemsens

Problem 2.1 (Original). *Suppose n countries collectively possess 1 large unit of currency in some distribution. Now suppose a random lottery system occurs repeatedly, randomly and uniformly selecting one country to receive a fixed fraction of all other countries' wealth. Is there a standard distribution describing how the wealth is distributed between the countries in the limit of number of countries and number of iterations?*

A project set to unravel the mathematics behind this problem was initiated by Prof. Misha Guysinsky and undergraduate student Maria Burago in early 2007. The problem they considered was as follows.

Problem 2.2 (Burago). *Suppose initial vector V is chosen with the sum of its components being 1. Does the distribution of values within the obtained vector V^* converge to Benford's distribution when the vector V is pre-multiplied by a product of stochastic matrices that are randomly chosen from a finite, well-defined set, as long as the ratio of the number of matrices in the product to the length of vector V converges to infinity?*

Definition 2.3. *A collection of numbers are said to obey b-Benford's distribution if for a number in this collection, the first digit takes value d, such that $d \in \{1, ..., b-1\}$ (in base $b \geq 2$) with a probability proportional to*

$$\log_b (d+1) - \log_b d = \log_b 1 + \frac{1}{d},$$

i.e. exactly the distance between d and $d+1$ on a logarithmic scale.

Date: August 31, 2017.

In her thesis, Burago details the precise stochastic matrices that correspond to the iterations acting on $V \in \mathbb{R}^n$. Suppose upon each iteration the chosen component $k \in [1, n]$ receives $\frac{1}{f}$ of all remaining wealth, paid for by the remaining countries in proportion to their wealth. Label such a left action on V as

$$A_k = \begin{pmatrix} \frac{f-1}{f} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{f-1}{f} & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{f} & \frac{1}{f} & 1 & \frac{1}{f} & \cdots & \frac{1}{f} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{f-1}{f} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{f-1}{f} \end{pmatrix},$$

$$V' = A_k V.$$

Models for this problem naturally arose to use as approximations. A new problem was forged:

Problem 2.4 (New)**.** *Suppose N points are distributed along a circle of circumference 1, and suppose our process is to randomly delete one point, and then to replace it by another point placed exactly at the next integer multiple modulo 1 of an irrational number. Show that in the limit of number of points and number of iterations that this set is equidistributed on the interval $[0, 1]$.*

Definition 2.5. *A sequence $(s_1, s_2, ...)$ of real number is said to be equidistributed on a non-degenerate interval $[a, b]$ if for any subinterval $[c, d]$ of $[a, b]$ we have*

$$\lim_{n\to\infty} \frac{|\{s_1, ..., s_n\} \cap [c, d]|}{n} = \frac{d-c}{b-a}.$$

Rather than prove the result directly, we wish to construct a model that we can prove provides an upper bound on the probabilities of cases greater than the n-Benford distribution. This model looks as follows.

Definition 2.6 (Model)**.** *Let N points be placed randomly along a circle of unit circumference. Then let each iteration performed be composed of a removal of $y \approx 1000$ points and then a replacement by sy points to our interval of length s and the remainder of points to the rest of the interval.*

Prof. Guysinsky and I tried many toy examples and small calculations, and here is tangent into these calculations. Suppose our process on N points involves an iteration of removing two points chosen uniform randomly from the N, and then replacing exactly 1 into our segment and another into the remainder of the interval. Suppose we begin with 0 points in our segment. Then we certainly will end this iteration with 1 point in our interval. To set up some notation, let $\mathcal{T} = (\mathcal{T}_{ij})$, where $\mathcal{T}_{ij} = \Pr(j \to i)$, or the probability of ending with i points in our segment after beginning with j. Obviously, $\mathcal{T}_{ij} = 0$ when $i - j > 1$. Let's perform some preliminary calculations.

- $\mathcal{T}_{00} = 0$,
- $\mathcal{T}_{0j} = 0$,
- $\mathcal{T}_{10} = 1$,
- $\mathcal{T}_{11} = \frac{1}{N} + \frac{N-1}{N}\frac{1}{N-1} = \frac{2}{N}$,
- $\mathcal{T}_{12} = \frac{2}{N}\frac{1}{N-1} = \frac{2}{N(N-1)}$,

- $\mathcal{T}_{21} = \frac{N-1}{N}\frac{N-2}{N-1} = \frac{N-2}{N}$,
- $\mathcal{T}_{(N-1)(N-1)} = \frac{2}{N}$
- $\mathcal{T}_{N(N-1)} = 0$,
- $\mathcal{T}_{(N-1)N} = 1$,
- $\mathcal{T}_{NN} = 0$.

Outside of these special cases, if $j > 1$, then $\mathcal{T}_{(j-1)j} = \Pr(j \to j-1) = \frac{j(j-1)}{N(N-1)}$; if $j > 0$, then $\mathcal{T}_{jj} = \Pr(j \to j) = 2\frac{j(N-j)}{N(N-1)}$; and if $j < N$, then $\mathcal{T}_{(j+1)j} = \Pr(j \to j+1) = \frac{(N-j)(N-j-1)}{N(N-1)}$. Therefore, our matrix $\mathcal{T}$ looks like

$$\mathcal{T} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 \\ 1 & \frac{2}{N} & \frac{2}{N(N-1)} & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{N-2}{N} & 4\frac{N-2}{N(N-1)} & \frac{(N-2)(N-3)}{N(N-1)} & \cdots & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{j(j-1)}{N(N-1)} & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 2\frac{j(N-j)}{N(N-1)} & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \frac{(N-j)(N-j-1)}{N(N-1)} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & \frac{2}{N(N-1)} & \frac{2}{N} & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Notice the sum of the columns must be 1. Our main goal is to show that this tridiagonal matrix. Label with $(a_n) \equiv (\mathcal{T}_{n(n-1)})_{n=1}^{N}$, the lower diagonal, with $(b_n) \equiv (\mathcal{T}_{nn})_{n=0}^{N}$ the main diagonal, and with $(c_n) \equiv (\mathcal{T}_{n(n+1)})_{n=0}^{N-1}$ the upper diagonal. From our relations, we have

$$a_n = \frac{(N-n)(N-n+1)}{N(N-1)};$$

$$b_n = \begin{cases} 0 & \text{if } n = 0, \\ 2\frac{n(N-n)}{N(N-1)} & \text{if } n > 0; \end{cases}$$

$$c_n = \begin{cases} 0 & \text{if } n = 0, \\ \frac{n(n+1)}{N(N-1)} & \text{if } n > 0. \end{cases}$$

Our goal is to show that the components of the eigenvector of this matrix $\mathcal{T}$ satisfy exponential relations up to some value, such as the approximately stable point $N/2$. For example, we see that the overwhelming trend is for the probability to increase when the number of points in our segment is below $N/2$, but we are only interested in the case of when we have too many points (more than $N/2$). We claim that past $x_{N/2}$, the following relation is satisfied for some fixed $d > 1$:

$$dx_{k+1} \leq x_k.$$

We also assume the approximations that when $N \approx n$, then $0 \leq a << b << c$, but $a_k + b_k + c_k \approx 1$. We suppose the exponential relation holds and investigate what constraints it leaves on our d. Notice from a general tridiagonal matrix, row N gives

$$a_N x_{N-1} + b_N x_N = x_N \implies x_{N-1} = x_N \cdot \frac{1 - b_N}{a_N}.$$

This condition along with our assumption of an exponential decaying sequence of eigenvector components implies $d \leq \frac{1-b_N}{a_N}$.
Generally, we begin with row $k > N/2$

$$a_k x_{k-1} + b_k x_k + c_k x_{k+1} = x_k,$$

and solving for x_{k-1}:

$$x_{k-1} = \frac{1}{a_k}\left[(1-b_k)x_k - c_k x_{k+1}\right] \geq d x_k,$$

so $x_k \leq \frac{c_k}{1-b_k-a_k d} x_{k+1}$. We will choose our d so that it is less than this coefficient but greater than 1.

$$d \geq \frac{c_k}{1-b_k-a_k d}$$
$$\Leftrightarrow 0 \geq d^2 - \frac{1-b_k}{a_k} d + \frac{c_k}{a_k}.$$

Recall that $1 - a_k - b_k \approx c_k$, and let $L = \frac{c_k}{a_k}$. Then we have

$$0 \geq d^2 - (L+1)d + L = (d-L)(d-1).$$

Notice $L = \frac{c_k}{a_k} >> 1$. Because $k \in (N/2, N]$ as an integer, take $d^* = \frac{1}{2} + \frac{1}{2} \min L_k = \frac{1}{2} + \min \frac{c_k}{2a_k}$. Such a quantity is only well defined if we impose the extra condition that $L_k = c_k/a_k > 1$ for all $k > N/2$. This d^* satisfies equations for all d_k and therefore admits $d^* x_{k+1} \leq x_k$, showing the distribution of eigenvector components is exponentially decaying.

Theorem 2.7. *The last half of the components of the eigenvector of a tridiagonal matrix $(a_n, b_n, c_n) \in M_{N\times N}(\mathbb{R})$ are exponentially decaying if $n \approx N$ implies $0 \leq a_n << b_n \approx c_n$ and that $1 \approx a_n + b_n + c_n$.*

The approximations like $1 \approx a_n + b_n + c_n$ refer to the following property

$$\lim_{N\to\infty} a_n + b_n + c_n = 1 \text{ uniformly, as } n \approx N.$$

For Burago's calculations, we could construct the same tridiagonal matrix and label it with a_n, b_n, c_n. Let $\mathcal{B} = (\mathcal{B}_{ij})$ where $\mathcal{B}_{ij} = \Pr(j \to i)$. The process of decreasing in general has probability $\mathcal{B}_{(j-1)j} = \Pr(j \to j-1) = \frac{j(m-1)}{mN}$ and is valid for $j > 0$; the probability of remaining constant is $\mathcal{B}_{jj} = \Pr(j \to j) = \frac{j+(N-j)(m-1)}{mN}$ and is valid for all j; and the probability of increasing is $\mathcal{B}_{(j+1)j} = \Pr(j \to j+1) = \frac{N-j}{mN}$ and is valid for $j < N$.

$$a_n = \frac{N-n+1}{mN};$$
$$b_n = \frac{n+(N-n)(m-1)}{mN};$$
$$c_n = \frac{(n+1)(m-1)}{mN}.$$

Do Burago's entries satisfy $0 \leq a << b << c$ when $n \approx N$? Actually, they are such that $0 \leq a << b \approx c$. This is alright, because in the previous general calculations we never invoked $b << c$, just that $a << c$. Certainly $a_k + b_k + c_k \approx 1$ as we wished. Therefore, we no longer need the explicit calculations of $P_0, P_1, ...$, etc. We have showed that P_i satisfy exponential relations, and Burago's Benford result is a consequence.

3. My Attempt at Skipping Linear Algebra Class

I want to generate a way to solve tridiagonal matrices. No explicit solution is possible, but we can at least create an inductive process to solving a tridiagonal matrix.

Theorem 3.1. *Let $X^{(n)} = (a_k, b_k, c_k)_{k=1}^n$ with augmented matrix having all 1's in the final column. The solution $x_i = 1/\beta_i$, where*

$$\beta_n = \frac{b_n \beta_{n-1}}{\beta_{n-1} - a_n},$$

and $(\beta_1, ..., \beta_{n-1})$ is the solution to $X^{(n-1)} = (a_k, b_k, c_k)_{k=1}^{n-1}$ with b_{n-1} replaced by

$$b'_{n-1} = \frac{b_{n-1}b_n - a_n c_{n-1}}{b_n - c_{n-1}}.$$

Proof. Checking that the formula works for some $X^{(2)}$ is easy. Begin with $X^{(n)} = (a_k, b_k, c_k)_{k=1}^n$ with all 1's on the augmented column. Call $X^{(n-1)} = (a_k, b_k, c_k)_{k=1}^{n-1}$ with b_{n-1} replaced with b'_{n-1} as above. Then solve $X^{(n-1)}$ inductively to be a diagonal matrix $(\beta_k)_{k=1}^{n-1}$. Then we replace in $X^{(n)}$ as follows

$$X^{(n)} = \begin{pmatrix} X^{(n-1)} & 0 \\ 0, ..., a_n & b_n \end{pmatrix}.$$

Subtracting the second to last row proportionally from the bottom row, we then rescale the row so 1 is left in the augmented row, showing

$$\beta_n = \frac{b_n \beta_{n-1}}{\beta_{n-1} - a_n}.$$

□

Now I will attempt to generalize our approach to proving the exponentially decreasing property of last half of the components of the eigenvector of some pent-diagonal matrix $X^{(n)} = (a_k, b_k, c_k, d_k, e_k)_{k=1}^n$. We wish to show that there exists a $\mu > 1$ such that $k > N/2$ implies $x_k \geq \mu x_{k+1}$. We generally have the recurrence relation

$$x_k = a_k x_{k-2} + b_k x_{k-1} + c_k x_k + d_k x_{k+1} + e_k x_{k+2}.$$

After repeatedly solving for the x_i with smallest index and invoking $x_i \geq \mu x_{i+1}$, then we obtain

$$0 \geq a_k \mu^4 + b_k \mu^3 + (c_k - 1)\mu^2 + d_k \mu + e_k \equiv \mathcal{P}(\mu).$$

The $(c_k - 1)$ term occurs because in the recurrence relation, we subtract x_k from both sides to decrement the coefficient on x_k by 1. We will consider the possible conditions that may arise, such as $a_k + b_k + c_k + d_k + e_k \approx 1$ when $k \approx n$, and perhaps $a_k << b_k << c_k << d_k << e_k$. Our previous theorem forces $b_k << c_k$ and $c_k \approx d_k$ at least. Using the first condition, we can barely simplify the problem to

$$\mu^4 + \frac{b_k}{a_k}(\mu^3 - 1) + \frac{c_k - 1}{a_k}(\mu^2 - 1) + \frac{d_k}{a_k}(\mu - 1) + \frac{1}{a_k} \leq 0.$$

However, we want to eventually factor this quadratic $\mathcal{P}$. Without invoking the general solution in radicals, I can make little progress. For example, I can prove the following.

Proposition 3.2. $\mathcal{P}(\mu) \mid (\mu - 1)$.

Proof. Assume $\mathcal{P}(\mu) = (\mu-1)(\alpha\mu^3+\beta\mu^2+\gamma\mu+\delta)$ for some $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Obviously, $\alpha = a_k$ and $\delta = -e_k$. Then the coefficient on μ^3 must satisfy

$$b_k = -\alpha + \beta \implies \beta = a_k + b_k.$$

The coefficient on μ must satisfy

$$d_k = -\gamma + \delta \implies \gamma = -d_k - e_k.$$

Finally, we verify this solution works for the coefficient on μ^2:

$$-\beta + \gamma = -a_k - b_k - d_k - e_k \approx c_k - 1.$$

By our estimate, this result holds for $k \approx n$. Thus,

$$\mathcal{P}(\mu) = (\mu - 1)(a_k\mu^3 + (a_k + b_k)\mu^2 - (d_k + e_k)\mu - e_k).$$

□

I have tried testing divisibility by some L analogous to that from the tridiagonal case, but I cannot find a working one. Obviously, higher diagonal matrices will require solving polynomials of degree one less than their number of nonzero diagonals.

4. Generalized Calculations for Wealth Transfer

To generalize Burago's calculations, I consider moving from value $i-2$ to i (in the general case). Although I don't agree with her endpoint calculations, but I can at least produce the same calculations for the general terms. below is a summary of the results.

$$P(i-2 \to i) = \frac{(n-i+2)(n-i+1)}{n(n-1)m^2}.$$

$$P(i-1 \to i) = \frac{(n-i+1)[2(n-1)(m-1) + (i-1)]}{n(n-1)m^2}.$$

$$P(i \to i) = \frac{(n-1)(n-i-1)(m-1)^2 + 4i(n-i)(m-1) + i(i-1))}{n(n-1)m^2}.$$

$$P(i+1 \to i) = \frac{2(i+1)(m-1)[i + (n-i+1)(m-1)]}{n(n-1)m^2}.$$

$$P(i+2 \to i) = \frac{(i+2)(i+1)(m-1)^2}{n(n-1)m^2}.$$

I'm sure I could probably generalize this procedure for more large diagonal matrices. Not. Enough. Time.