MATH 5763 Homework 3 Due Fri, 2/18/11

Problem 1.

(a) Consider two urns A and B. Initially the urn A contains N black balls and the urn B contains N white balls. At each step, one ball is selected at random from each urn and the two balls interchange. Let X_n denote the number of white balls in the urn A at time n. Determine the transition matrix $\mathbf{P}$.

Remark: The elements p_{0j}, p_{Nj}, and p_{ij} for $1 \le i \le N-1$ should be treated separately. One way to check your results is to verify that $\mathbf{P}$ is a stochastic matrix.

(b) Consider two urns A and B containing a total N balls together. At each time, a ball is selected at random (all selections are equally likely) from among the totality of N balls. Then an urn is selected at random: urn A is selected with probability p and urn B is selected with probability $1-p$. And the ball previous drawn is placed in this urn. Let X_n denote the number of balls in A at time n. Determine the transition matrix $\mathbf{P}$.

Problem 2. The newspaper is delivered every morning to Brandon's house. Brandon reads the newspaper at 8 a.m., and puts it on a pile after reading it. However, if the pile contains 5 newspapers after he puts the newspaper on it, he throws all the newspapers in the pile (including the new one) in the recycle bin. Also, at 6 p.m. every evening, with probability $\frac{1}{3}$, Brandon takes all the papers in the pile and puts them in the recycle bin. Model this by a Markov chain and write the transition matrix.

Hint: Let X_n be the number of papers at 6:01 p.m. on day n.

Problem 3. Consider a Markov chain whose state space consists of five states: α, β, γ, δ, ϵ, and whose transition matrix is the following:

$$\mathbf{P} = \begin{array}{c} \begin{matrix} \alpha & \beta & \gamma & \delta & \epsilon \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \begin{matrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \epsilon \end{matrix}$$

(a) Draw a diagram with arrows (where each arrow from state i to state j represents a nonzero probability p_{ij}), and identify the transient and the recurrent states (do not do any computations yet). You will find that two states are transient (denote the set of transient states by D), and there will be two closed and irreducible sets of recurrent states (one of them – call it C_1 – will consist of two states, and the other will consist of only one state – call this set C_2).

(b) Now relabel the states α, β, γ, δ, ϵ as 1, 2, 3, 4, 5, in such a way that the $C_1 = \{1, 2\}$, $C_2 = 3$, and the states 4 and 5 to be the transient states, i.e., $D = \{4, 5\}$. In C_1, let state 1 be the state with one-step probability for transition to itself equal to $\frac{1}{3}$; in D, let state 4 be the state with nonzero one-step probability for transition to itself.

(c) Carefully write the one-step transition probability matrix $\widetilde{\mathbf{P}}$ with the relabeled states. It should look like this:

$$\widetilde{\mathbf{P}} \;=\; \left(\begin{array}{c|c|c} \mathbf{C}_1 & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_2 & \mathbf{0} \\ \hline * & * & \mathbf{T} \end{array}\right) ,$$

where $\mathbf{0}$ are matrices (of appropriate size) with all entries equal to zero, while the stars represent matrices that are generally not zero (but nothing more concrete can be said about them in general).

Check that $\mathbf{C}_1$ and $\mathbf{C}_2$ are stochastic matrices, while $\mathbf{T}$ is *not* a stochastic matrix.

Problem 4. Let V be a d-dimensional linear space (e.g., $\mathbb{R}^d$), and $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_d$ be a basis in V; in this basis every vector $\mathbf{u} \in V$ can be written as $\mathbf{u} = \sum_{i=1}^{d} u_i\, \mathbf{e}_i$. Let $\mathsf{P} : V \to V$ be a linear transformation (sometimes linear transformations are called "linear operators"). The matrix $\mathbf{P} = (p_{ij})$ of the linear transformation P in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_d$ is defined by

$$\mathsf{P}\mathbf{e}_j =: \sum_{i=1}^{d} p_{ij}\, \mathbf{e}_i \ ,$$

i.e., $p_{ij} = (\mathsf{P}\mathbf{e}_j)_i$ is the ith component of the vector $\mathsf{P}\mathbf{e}_j$ in the basis $\mathbf{e}_i$. If one changes the basis from $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_d$ to $\widetilde{\mathbf{e}}_1$, $\widetilde{\mathbf{e}}_2$, …, $\widetilde{\mathbf{e}}_d$, then the matrix of the linear transformation P changes from $\mathbf{P}$ to $\widetilde{\mathbf{P}} = (\widetilde{p}_{ij})$. Suppose that the "new" basis $\widetilde{\mathbf{e}}_i$ is expressed in terms of the "old" basis $\mathbf{e}_i$ as

$$\widetilde{\mathbf{e}}_j = \sum_{i=1}^{d} c_{ij}\, \mathbf{e}_i \ ,$$

and the matrix $\mathbf{C}$ is defined by $\mathbf{C} = (c_{ij})$. Clearly, the entries of $\mathbf{C}$ are the components of the new basis vectors in the old basis – indeed, comparing the formula above with $\widetilde{\mathbf{e}}_j = \sum_{i=1}^{d} \big(\widetilde{\mathbf{e}}_j\big)_i \mathbf{e}_i$, we see that $c_{ij} = \big(\widetilde{\mathbf{e}}_j\big)_i$. In other words, the jth column of the matrix $\mathbf{C}$ is the vector $\widetilde{\mathbf{e}}_j$ (written in the old basis $\mathbf{e}_i$):

$$\mathbf{C} = (\, \widetilde{\mathbf{e}}_1 \,|\, \widetilde{\mathbf{e}}_2 \,|\, \cdots \,|\, \widetilde{\mathbf{e}}_d) \ .$$

It is easy to see that

$$\mathbf{e}_j = \sum_{i=1}^{d} d_{ij}\, \widetilde{\mathbf{e}}_i \ ,$$

where $(d_{ij}) = \mathbf{C}^{-1}$ is the inverse matrix of $\mathbf{C}$. In the new basis the linear transformation P has matrix $\widetilde{\mathbf{P}} = (\widetilde{p}_{ij})$, where $\widetilde{\mathbf{P}} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$:

$$\begin{aligned} \sum_m \widetilde{p}_{mj}\, \widetilde{\mathbf{e}}_m &= \mathsf{P}\widetilde{\mathbf{e}}_j = \mathsf{P}\left(\sum_i c_{ij}\, \mathbf{e}_i\right) = \sum_i c_{ij}\, \mathsf{P}\mathbf{e}_i = \sum_i d_{ij} \sum_k p_{ki}\, \mathbf{e}_k \\ &= \sum_i c_{ij} \sum_k p_{ki} \sum_m d_{mk}\, \widetilde{\mathbf{e}}_m = \sum_m \left(\sum_i \sum_k d_{mk}\, p_{ki}\, c_{ij}\right) \widetilde{\mathbf{e}}_m \ . \end{aligned}$$

If the matrix $\mathbf{P}$ has d distinct real eigenvalues λ_j, and the corresponding eigenvectors are $\widetilde{\mathbf{e}}_j$, then in the basis $\widetilde{\mathbf{e}}_j$, the matrix $\widetilde{\mathbf{P}}$ of the linear transformation P will be diagonal (with the eigenvalues λ_j on the diagonal).

In this problem you will apply these techniques in order to find high powers of stochastic matrices.

(a) Consider the stochastic matrix

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

The eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2}$, and $\lambda_3 = -\frac{1}{6}$, and the corresponding eigenvectors are

$$\widetilde{\mathbf{e}}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ , \quad \widetilde{\mathbf{e}}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \ , \quad \widetilde{\mathbf{e}}_3 = \begin{pmatrix} 3 \\ -4 \\ 3 \end{pmatrix} \ .$$

The matrix $\mathbf{C}$ and its inverse are

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 0 & -4 \\ 1 & 1 & 3 \end{pmatrix} \ , \qquad \mathbf{C}^{-1} = \frac{1}{14} \begin{pmatrix} 4 & 6 & 4 \\ -7 & 0 & 7 \\ 1 & -2 & 1 \end{pmatrix} \ .$$

What is the matrix $\widetilde{\mathbf{P}} = \mathbf{C}^{-1}\mathbf{P}\mathbf{C}$?

Remark: I used Mathematica to do some of the calculations. To enter the matrix $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ in Mathematica, type `B={{1,2},{3,4}}` and then hold down the SHIFT key and press ENTER. To find its inverse of $\mathbf{B}$, type `Inverse[B]` and again press press ENTER while holding down the SHIFT key.

(b) Show that, if $\widetilde{\mathbf{P}}$ is an arbitrary square matrix, $\mathbf{C}$ is an invertible matrix of the same size as $\widetilde{\mathbf{P}}$, and $\mathbf{P} = \mathbf{C}\widetilde{\mathbf{P}}\mathbf{C}^{-1}$, then $\mathbf{P}^n = \mathbf{C}\widetilde{\mathbf{P}}^n\mathbf{C}^{-1}$ for any $n \in \mathbb{N}$.

(c) Use your result of part (b) to compute $\mathbf{P}^n$ for any $n \in \mathbb{N}$, where $\mathbf{P}$ is the stochastic matrix from part (a).

(d) Let X_n be a discrete time, discrete state space Markov chain whose state space consists of three states (labeled 1, 2, 3) and whose transition probabilities are given by the stochastic matrix $\mathbf{P}$ from part (a), i.e., $\mathbf{P}(X_{n+1} = j | X_n = i) = p_{ij}$. Let $\mathbf{a} = (a_1\ a_2\ a_3)$ be the initial probability distribution, i.e., $\mathbb{P}(X_0 = i) = a_i$. Find $\mathbb{P}(X_n = i)$ and $\lim_{n\to\infty} \mathbb{P}(X_n = i)$.

Problem 5. A *linear homogeneous recurrence relation of order d with constant coefficients* is an equation of the form

$$x_n = b_1 x_{n-1} + b_2 x_{n-2} + \cdots + b_d x_{n-d} \ , \tag{1}$$

where the d coefficients b_1, b_2, …, b_d are constants. Solving such relations is very similar to solving linear homogeneous ordinary differential equations of order d with constant coefficients. Similarly to the case of differential equations, one talks about a *general solution* of (1) (the general solution contains d arbitrary constants), and for the *particular solution* of (1), which satisfies not only (1), but also *initial conditions*,

$$x_0 = a_0 \ , \ x_1 = a_1 \ , \ \ldots \ , \ x_{d-1} = a_{d-1} \ . \tag{2}$$

To find the general solution of (1), set $x_j = \lambda^j$ in the equation to obtain (after dividing by λ^{n-d}) the equation

$$P(\lambda) = \lambda^d - b_1\lambda^{d-1} - b_2\lambda^{d-2} - \cdots - b_d = 0 \ ,$$

called the *characteristic equation* of (1) (the polynomial P is called the *characteristic polynomial* of (1)). Find all roots λ_j of the characteristic equation.

If all the roots of the are real and distinct (i.e., the roots are λ_1, … λ_d with $\lambda_j \in \mathbb{R}$ and $\lambda_i \neq \lambda_j$ for $i \neq j$), then the general solution of (1) has the form

$$x_n = C_1\lambda_1^n + C_2\lambda_2^n + \cdots + C_d\lambda_d^n \ ,$$

where C_1, C_2, …, C_d are arbitrary constants. To find the particular solutions that satisfies both the relation (1) and the initial condition (2), one needs to express the constants C_1, C_2, …, C_d through the initial conditions a_0, a_1, …, a_{d-1}.

If $\lambda_j \in \mathbb{R}$ is a real root of the characteristic equation with multiplicity s_j (i.e., if the characteristic polynomial contains the factor $(\lambda - \lambda_j)^{s_j}$), then the corresponding term in the general solution x_n of the recurrence relation (1) is

$$\left(C_1^{(j)} + C_2^{(j)} n + C_3^{(j)} n^2 + \cdots C_{s_j}^{(j)} n^{s_j - 1} \right) \lambda_j^n \ ,$$

where $C_k^{(j)}$ are arbitrary constants. For example, if $P(\lambda) = (\lambda - 5)\lambda^2(\lambda + 7)^3$, then the general solution of the corresponding recurrence relation is

$$x_n = C_1 5^n + C_2 + C_3 n + (C_4 + C_5 n + C_6 n^2)(-7)^n$$

(after relabeling the arbitrary constants).

(a) The Fibonacci numbers F_n are defined by

$$F_0 = 1 \ , \quad F_1 = 1 \ , \qquad F_n = F_{n-1} + F_{n-2} \ \text{ for } n \geq 2 \ .$$

First few FIbonacci numbers are 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ... Find an explicit formula for F_n.

(b) Let $p \in (0, 1)$ be a constant. Find all functions that satisfy the relation

$$f(n) = (1 - p)f(n - 1) + pf(n + 1) \ .$$

You will have to consider the cases $p \neq \frac{1}{2}$ and $p = \frac{1}{2}$ separately.