CS4220 Assignment 3 Due: 2/21/13 (Thur) at 11pm

You must work either on your own or with one partner. You may discuss background issues and general solution strategies with others, but the solutions you submit must be the work of just you (and your partner). If you work with a partner, you and your partner must first register as a group in CMS and then submit your work as a group. Each problem is worth 5 points. One point may be deducted for poor style.

Topics: stochastic matrix, diagonal dominance, null space, LU, Cholesky, LDL, inertia

1 Stationary Vector of a Structured Stochastic Matrix

A matrix $T \in \mathbb{R}^{n\times n}$ is *stochastic* if its entries are nonnegative and the entries in each column sum to one. In short, it is a matrix of transition probabilities where t_{ij} is the probability of transitioning from state j to state i in one "time step." Suppose $x \in \mathbb{R}^n$ is such that x_j is the number of "objects" in state j. If $y = Tx$ then

$$y_i \;=\; \sum_{j=1}^{n} t_{ij}x_j$$

specifies the expected number of objects in state i after one time step. With a few extra reasonable assumptions it can be shown that there is a vector x with nonnegative entries so that $Tx = x$. Such a vector is called a *stationary vector* and their computation is of central importance in many stochastic modelling problems. Note that if x is a stationary vector then so is any positive multiple of x.

One way to compute a stationary vector is to compute the LU factorization of the singular matrix $A = T-I$. It can be shown that if

$$PA = LU$$

then $U(1{:}n-1, 1{:}n-1)$ is nonsingular and $U(n,n) = 0$, e.g.,

$$U \;=\; \left[\begin{array}{ccc|c} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ \hline 0 & 0 & 0 & 0 \end{array}\right].$$

Thus,

$$Tx = x \;\Rightarrow\; Ax = (T-I)x = 0 \;\Rightarrow\; PAx = 0 \;\Rightarrow\; LUx = 0 \;\Rightarrow\; Ux = 0$$

and so

$$\left[\begin{array}{c|c} U(1{:}n-1, 1{:}n-1) & U(1{:}n-1, n) \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c} x(1{:}n-1) \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 0 \end{array}\right]$$

(Since we are looking null vectors, there is no harm in setting $x(n) = 1$.) By examining the "top half" of this equation we see that

$$x(1{:}n-1) = -U(1{:}n-1, 1{:}n-1)\backslash U(1{:}n-1, n) \tag{1}$$

In this problem you use this solution framework on a stochastic matrix T that has an "almost" tridiagonal structure which we next describe.

The starting point in the derivation of our specially structured problem is to assume that we are given n angles

$$0 \le \theta_1 < \theta_2 < \cdots < \theta_n < 2\pi.$$

Let Q_i be the point $(\cos(\theta_i), \sin(\theta_i))$, $i = 1{:}n$. Define the squared-distance reciprocals $r_1, \ldots, r_n$ by

$$\frac{1}{r_i} \;=\; \begin{cases} \text{Squared Euclidean distance between } Q_i \text{ and } Q_{i+1} \text{ if } 1 \le i \le n-1 \\[2ex] \text{Squared Euclidean distance between } Q_i \text{ and } Q_1 \text{ if } i = n \end{cases}$$

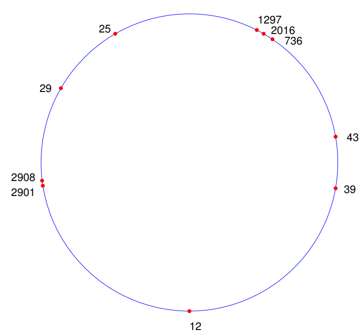
The following $n = 5$ example fully defines the "general n" version of T:

$$T = \begin{bmatrix} 2 & r_1 & 0 & 0 & r_5 \\ r_1 & 2 & r_2 & 0 & 0 \\ 0 & r_2 & 2 & r_3 & 0 \\ 0 & 0 & r_3 & 2 & r_4 \\ r_5 & 0 & 0 & r_4 & 2 \end{bmatrix} D^{-1} \qquad D = \begin{bmatrix} 2+r_5+r_1 & 0 & 0 & 0 & 0 \\ 0 & 2+r_1+r_2 & 00 & 0 & \\ 0 & 0 & 2+r_2+r_3 & 0 & \\ 0 & 0 & 0 & 2+r_3+r_4 & 0 \\ 0 & 0 & 0 & 0 & 2+r_4+r_5 \end{bmatrix}$$

You should confirm that T is stochastic and that $T - I$ is diagonally dominant. Stochastic matrices that are built this way from the θ values will be called *circular stochastic matrices*.

Here is an $n = 10$ visualization of the underlying stochastic process together with a display of the stationary vector

$$x^T \;=\; [\; 43 \;\; 736 \;\; 2016 \;\; 1297 \;\; 25 \;\; 29 \;\; 2908 \;\; 2901 \;\; 12 \;\; 39 \;]$$



(The stationary vector has been normalized so that the sum of its components is approximately 10000.) In the schematic, think of the dots as islands and that at each time step the inhabitants either stay put or hop to an "adjacent" island. We provide a little intuition. If you are on an island that is very close to exactly one other island, then the probability is very high that you will hop there. (Think about how the reciprocal distances define the transition probabilities. If island j is such an island, look at the three probabilities in $T(:, j)$.) If you are in a tight cluster of islands, it will be "pretty hard" to hop outside the cluster. These observations are confirmed by the components of the stationary vector. Observe that the bigger values in the stationary vector are associated with island clusters.

Complete the following two functions so that they perform as specified:

```
  function T = MakeTransition(theta)
% theta is a column n-vector with the property that 0 <= theta(1) < ... < theta(n) < 2*pi
% T is the associated circular stochastic matrix.

  function x = Stationary(T)
% T is an nxn circular stochastic matrix
% x satisfies Tx = x and its entries are nonnegative and sum to one.
```

Regarding the implementation of `Stationary`, you are allowed to use `lu` to compute $P(T - I) = LU$. Experiment and draw some conclusions about the structure of U. For full credit, your method for solving (1) must exploit this structure. (Do NOT use \.) Submit `MakeTransition` and `Stationary` to CMS. A test script `P1` is available on the website.

2 A Structured Indefinite System Solver

Write a function `[y,z] = SolveSym(A,C,f,g)` that solves the symmetric indefinite linear system

$$\begin{bmatrix} A & C \\ C^T & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite, $C \in \mathbb{R}^{n \times r}$ has full column rank, $f \in \mathbb{R}^n$, and $g \in \mathbb{R}^r$. (The full rank assumption guarantees that the overall system is nonsingular.) We can derive a relevant structured LU factorization for this problem by equating blocks in

$$\begin{bmatrix} A & C \\ C^T & 0 \end{bmatrix} = LU \equiv \begin{bmatrix} G & 0 \\ H & K \end{bmatrix} \begin{bmatrix} G & 0 \\ H & -K \end{bmatrix}^T, \tag{2}$$

i.e.,

- (1,1): $A = GG^T$ so get G via `chol(A,'lower')`.
- (2,1): $C^T = HG^T$ so get H by solving $GH^T = C$.
- (2,2): $0 = HH^T - KK^T$ so get K by another Cholesky.

You are allowed to use `chol` and you can use \ to solve triangular systems. Submit `SolveSym` to CMS. It should include a brief comment that explains the output when the test script `P2` is executed.

3 The Inertia of a Symmetric Matrix

If $A \in \mathbb{R}^{n \times n}$ is symmetric, then its *inertia* is a triplet of integers $(nNeg, nZero, nPos)$ which are respectively the number of negative, zero, and positive eigenvalues. One way to compute the inertia is to invoke the MATLAB function `schur` which computes the *Schur decomposition*:

$$Q^T A Q = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n), \qquad Q^T Q = I.$$

To get Q and Λ just execute `[Q,Lambda] = schur(A)`. If we want just the inertia and not the actual eigenvalues, then using `schur` is overkill and expensive. In this problem you exploit the *Sylvester Law of Inertia* and make use of the MATLAB function `ldl`. Sylvester's Law states that if $A \in \mathbb{R}^{n \times n}$ is symmetric and $X \in \mathbb{R}^{n \times n}$ is nonsingular, then A and $X^T AX$ have the same inertia. (Note, this is not a similarity transformation–the theorem does not require X to be orthogonal.) We show how to use `ldl` to "reveal" A's inertia. The call `[L,D,P] = ldl(A)` computes the factorization

$$P^T AP = LDL^T$$

where P is a permutation, L is unit lower triangular, and D is block diagonal. If $X^T = L^{-1}P^T$ then A and $X^T AX = D$ have the same inertia. This is progress because it is easy to count the positive and negative eigenvalues of D since the diagonal blocks are so small in dimension. For example, if

$$D = \begin{bmatrix} d_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & d_{22} & d_{23} & 0 & 0 & 0 \\ 0 & d_{32} & d_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & d_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & d_{55} & d_{66} \\ 0 & 0 & 0 & 0 & d_{65} & d_{66} \end{bmatrix}$$

then the eigenvalues of D are d_{11}, and d_{44} together with the eigenvalues of the symmetric matrices

$$D_2 = \begin{bmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{bmatrix} \qquad \text{and} \qquad D_5 = \begin{bmatrix} d_{55} & d_{56} \\ d_{65} & d_{66} \end{bmatrix}$$

You can use `schur` to get the eigenvalues of these 2x2 diagonal "bumps." Implement the following function:

```
  function [nNeg,nPos] = Sylvester(A)
% A is nonsingular and symmetric.
% nNeg and nPos are the number of negative and positive eigenvalues.
```

Submit your implementation to CMS. A test script `P3` is available on the course website.