Day 3 Contest: problem D

Moscow Pre-finals Workshop 2020

April 22, 2020

1 Notation

A is an $n \times m$ matrix if it has n rows and m columns. A_{ij} denotes the element in the i-th row and j-th column. Throughout, all indices are 0-based. If A is a vector ($n \times 1$ matrix) the second index is omitted ($A_i = A_{i,0}$).

When i is an index in the sum of matrix elements, the range is naturally assumed to be over the entire corresponding dimension.

The *scalar product* $\langle v, u \rangle$ of two vectors of equal length is equal to $\sum_i v_i \cdot u_i$.

I denotes the *identity matrix*: $I_{ij} = 1$ when $i = j$, and 0 otherwise. $[i]$ denotes the vector with element i equal to 1, and all other equal to 0. Dimensions of I and $[i]$ are usually clear from context, and generally omitted to avoid clutter.

$A + B$ is an element-wise sum of matrices A and B of equal dimensions. AB is the *matrix product*: $(AB)_{ij} = \sum_k A_{ik} B_{kj}$. For a square matrix A, A^{-1} is the *inverse matrix* (when it exists): $AA^{-1} = A^{-1}A = I$.

2 Absorbing Markov chains

Consider a random process with n states, at least one of which is *absorbing*. Let A be an $n \times n$ matrix, where A_{ij} is the probability to go from state i to state j. If i is absorbing, then all $A_{ij} = 0$, and if i is non-absorbing, we assume $\sum_j A_{ij} = 1$.

Consider vector E, where E_i is the expected number of steps to reach an absorbing state from the state i. By linearity of expectation we must have $E = AE + \mathbf{1}_n$, where $\mathbf{1}_n$ is the n-vector containing 1's in all non-absorbing states, and 0 in all absorbing states. It follows that $E = (I - A)^{-1}\mathbf{1}_n$, whenever $I - A$ has an inverse.

3 Eigenvectors and eigenvalues

If for a square matrix A and a non-zero vector v we have $Av = \lambda v$ for some real number λ, then v is called an *eigenvector*, and λ a corresponding *eigenvalue*.

A collection of n-vectors $v_0, \ldots, v_{n-1}$ (each of size n) is called *orthonormal* if $\langle v_i, v_j \rangle = 1$ when $i = j$, and 0 otherwise. For any vector x, we have $x = \sum_i \langle x, v_i \rangle v_i$.

An orthonormal collection $v_0, \ldots, v_{n-1}$ is an (orthonormal) *eigenbasis* of an $n \times n$ matrix A if all v_i are eigenvectors of A.

The following properties follow from definitions.

- Any orthonormal set of vectors $v_0, \ldots, v_{n-1}$ is an eigenbasis of I with all eigenvalues equal to 1.
- If $v_0, \ldots, v_{n-1}$ is an eigenbasis of A with eigenvalues λ_i, and also an eigenbasis of B with eigenvalues μ_i, then it is also:
  - (for any real x, y) an eigenbasis of $xA + yB$ with eigenvalues $x\lambda_i + y\mu_i$;
  - an eigenbasis of AB with eigenvalues $\lambda_i \mu_i$.
- A has an inverse A^{-1} when all its eigenvalues λ_i are non-zero. An eigenbasis $v_0, \ldots, v_{n-1}$ of A is then also an eigenbasis of A^{-1} with eigenvalues λ_i^{-1}.

4 1D random walk

Consider a random walk on a segment $[0, n]$. From an integer point x we equiprobably go to $x-1$ or $x+1$. Points 0 and n are absorbing.

The transition $(n+1) \times (n+1)$ matrix A is given by $A_{ij} = \frac{1}{2}$ when $|i-j| = 1$ and $i \neq 0, n$, and 0 otherwise.

For any $i = 1, \ldots, n-1$, define the $(n+1)$-vector v_i by $v_{ij} = \sin\left(\frac{ij\pi}{n}\right)$.

Statement.

- v_i is an eigenvector of A with eigenvalue $\lambda_i = \cos\left(\frac{i\pi}{n}\right)$.
- $\langle v_i, v_j \rangle = \frac{n}{2}$ when $i = j$, and 0 otherwise.

Proof. Trigonometry.

□

Corollary. Vectors $\sqrt{\frac{2}{n}} v_i$, together with vectors $v_0 = [0]$ and $v_n = [n]$, form an orthonormal eigenbasis. $\lambda_0 = \lambda_n = 0$.

It can be shown that the expectation vector E for this process can be given by $E_i = i(n-i)$. Still, let's find it with another method, as a simpler approximation for the full problem. Recall that $E = (I-A)^{-1}\mathbf{1}_{n+1}$. Let us represent $\mathbf{1}_{n+1} = \sum_i \langle v_i, \mathbf{1}_{n+1} \rangle v_i = \sum_i x_i v_i$.

Statement. $x_0 = x_n = 0$. $x_i = \sqrt{\frac{2}{n}} \sum_k \sin\left(\frac{ik\pi}{n}\right) = 0$ when i is even, and $\sqrt{\frac{2}{n}} \cot \frac{i\pi}{2n}$ when i is odd.

Proof. More trigonometry.

□

Since all $\lambda_i \neq 1$, $(I-A)^{-1}$ has the same eigenbasis v_i with eigenvalues $\frac{1}{1-\lambda_i}$. Thus

$$E = (I-A)^{-1}\mathbf{1}_{n+1} = \sum_k (I-A)^{-1} x_k v_k = \sum_k \frac{x_k v_k}{1-\lambda_k}.$$

Unrolling this for a single E_i, we get a nice expression

$$E_i = \sqrt{\frac{2}{n}} \sum_{k \text{ odd}} \frac{\sin \frac{ik\pi}{n} \cot \frac{k\pi}{2n}}{1 - \cos \frac{ik\pi}{n}}.$$

...which we know to be equal $i(n-i)$. Anyway...

5 Kronecker matrix product

As a matter of notation, if we know that indices i and j range in $[0, n)$ and $[0, m)$ respectively, let us denote $i \otimes j = i \cdot m + j$, so that $i \otimes j$ defines a bijection between pairs (i, j) and indices in range $[0, n \cdot m)$.
This may be confusing, but we will soon need matrices where a single dimension corresponds to all cells in a grid, and this is the simplest way to denote range simultaneously in the matrix and in the grid (that is, without introducing tensors...).

If A is an $a \times b$ matrix, and B is a $c \times d$ matrix, then $A \otimes B$ is the *Kronecker product* of A and B. $A \otimes B$ is a $(a \cdot c) \times (b \cdot d)$ matrix, defined by

$$(A \otimes B)_{i \otimes k, j \otimes l} = A_{ij} \times B_{kl}.$$

Informally, to obtain $A \otimes B$, construct it out of $a \times b$ block copies of matrix B, where the block (i, j) is multiplied by B_{ij}. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 5 \end{pmatrix} \otimes \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \\ 0 & 0 & 25 & 30 \\ 0 & 0 & 35 & 40 \end{pmatrix}.$$

Statement. Suppose that:

- A and B are square matrices;
- v_i is an (orthonormal) eigenbasis of A with eigenvalues λ_i;
- u_j is an (orthonormal) eigenbasis of B with eigenvalues μ_j.

Then $w_{i \otimes j} = v_i \otimes u_j$ is an (orthonormal) eigenbasis of $A \otimes B$ with eigenvalues $\nu_{i \otimes j} = \lambda_i \cdot \mu_j$.

Note that $v_i \otimes u_j$ is a Kronecker product of vectors as $n \times 1$ and $m \times 1$ matrices.

Statement. Suppose that vectors a, b (possibly of different sizes) satisfy $a = \sum_i x_i v_i$, $b = \sum_j y_j u_j$. Then $a \otimes b = \sum_{i,j} x_i y_j (v_i \otimes u_j)$.

6 2D random walk, without the missing cell

Consider a random walk on the grid $[0, n] \times [0, m]$, where transitions to the four adjacent points are equiprobable, and all border points are absorbing. In the remaining part i ranges in $[0, n + 1)$, and j ranges in $[0, m + 1)$. To refer to a point (i, j), or its index in a vector or matrix, we naturally use $i \otimes j \in [0, (n + 1) \cdot (m + 1))$.

Let A and B be transition matrices for the 1D random walks on $[0, n]$ and $[0, m]$ respectively. Denote v_i and u_j their orthonormal eigenbases produced in section 4, and λ_i and μ_j corresponding eigenvalues.

Statement. The transition matrix T for the random walk on $[0, n] \times [0, m]$ is equal to $\frac{1}{2}(A \otimes I + I \otimes B)$. *Note that the two I's here have different sizes.* It has an orthonormal eigenbasis $w_{i \otimes j} = v_i \otimes u_j$ with eigenvalues $\nu_{i \otimes j} = \frac{1}{2}(\lambda_i + \mu_j)$.

Similar to section 4, we now want to find $E = (I-T)^{-1}\mathbf{1}_{(n+1)(m+1)}$. Note that $\mathbf{1}_{(n+1)(m+1)} = \mathbf{1}_{n+1} \otimes \mathbf{1}_{m+1}$. Since we had representations $\mathbf{1}_{n+1} = \sum_i x_i v_i$, $\mathbf{1}_{m+1} = \sum_j y_j v_j$, we have that $\mathbf{1}_{(n+1)(m+1)} = \sum_{i,j} x_i y_j (v_i \otimes u_j)$ — a representation of $\mathbf{1}_{(n+1)(m+1)}$ in the basis $v_i \otimes u_j$.

We also know that $(I-T)^{-1}$ exists, and $(I-T)^{-1}(v_i \otimes u_j) = \frac{v_i \otimes u_j}{1-\frac{1}{2}(\lambda_i+\mu_j)}$. With this we can find

$$E = (I-T)^{-1}\mathbf{1}_{(n+1)(m+1)} = \sum_{i,j} \frac{x_i y_j (v_i \otimes u_j)}{1-\frac{1}{2}(\lambda_i+\mu_j)}.$$

A few notable things:

- if (i,j) is a border point, then x_i or y_j is zero, thus we can only sum over (i,j) inside the grid;
- if (i,j) is not a border point, all $x_i, y_j, \lambda_i, \mu_j$ and all components of v_i and u_j have simple trigonometric expressions given in section 4 (which we do not reproduce to avoid clutter);
- if we know all $v_i, u_j, x_i, y_j, \lambda_i, \mu_j$, a single component of E can be found in $O(nm)$ time.

7 2D random walk, with the missing cell

Let us now consider a random walk from the problem itself, where an extra point (i_a, j_a) is absorbing. We will still use the same transition matrix T to find the solution.

Let E' be the expectation vector for the modified process. It has to satisfy the following:

- $E'_{i_a \otimes j_a} = 0$;
- the linearity of expectation $E' = TE' + \mathbf{1}_{(n+1)(m+1)}$ has to be obeyed at all points, except (i_a, j_a). This means $E' = (I-T)^{-1}1(\mathbf{1}_{(n+1)(m+1)} + \delta[i_a \otimes j_a])$ for some real value of δ (recall that $[i_a \otimes j_a]$ is the vector with the only non-zero component $i_a \otimes j_a$ equal to 1).

Substituting the "unperturbed" answer from the previous section, we have $E' = E + \delta(I-T)^{-1}[i_a \otimes j_a]$. Since $E'_{i_a \otimes j_a}$ must be zero, we have

$$\delta = -\frac{E_{i_a \otimes j_a}}{((I-T)^{-1}[i_a \otimes j_a])_{i_a \otimes j_a}}.$$

The problem of finding $(I-T)^{-1}[i_a \otimes j_a]$ is similar to finding $E = (I-T)^{-1}\mathbf{1}_{(n+1)(m+1)}$. Since $[i_a \otimes j_a] = [i_a] \otimes [j_a]$ (with the factors being $(n+1)$- and $(m+1)$-vectors respectively), we can use this to represent $[i_a \otimes j_a]$ in the basis $v_i \otimes u_j$. In fact, $[i_a] = \sum_i v_{ii_a} v_i$, $[j_a] = \sum_j u_{jj_a} u_j$, thus $[i_a \otimes j_a] = \sum_{i,j} v_{ii_a} u_{jj_a} (v_i \otimes u_j)$. It follows that

$$(I-T)^{-1}[i_a \otimes j_a] = \sum_{i,j} \frac{v_{ii_a} u_{jj_a} (v_i \otimes u_j)}{1-\frac{1}{2}(\lambda_i+\mu_j)},$$

which we can use to find δ and finally obtain $E' = E + \delta(I - T)^{-1}[i_a \otimes j_a]$. Again, any single element of E' can be obtained in $O(nm)$, and most terms have simple trigonometric expressions found in section 4.