Advanced Scientific Computing: Numerical Methods AM205
Harvard University
Professor Chris Rycroft

December 1, 2017
Crypto project

Crypto project

Name: Yihang Yan, Tomas Gudmundsson, Nathaniel Stein

1

Jacobi method

In order to find eigenvalues and eigenvectors of the covariance matrix, A, we use the Jacobi method. The method works by finding the largest non-diagonal element at location (i, j) and makes it zero by doing a plane rotation on rows and columns i and j. The Jacobi method is quite efficient with quadratic convergence and can be parallelized easily.

The rotation for matrix A works as follows:

$$A' = P^T_{i,j,\theta} \cdot A \cdot P_{i,j,\theta} \tag{1}$$

where $P_{i,j,\theta} = P(i, j, \theta)$ is a Givens rotation matrix with the form:

$$P(i,j,\theta) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -s & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \tag{2}$$

This matrix has ones on the diagonal except where $c = cos(\theta)$ at locations (i,i) and (j,j). All other elements are 0 except $s = sin(\theta)$ at locations (i,j) and (j,i).

When the Givens rotation matrix, $P_{i,j,\theta}$, is multiplied with another matrix, A, as $P \cdot A$ it simulates a clockwise rotation in the plane by an angle θ in order to nullify the element at location (i, j) and only affects rows and columns i and j in the process.

In order to figure out the values of c, s and θ we will simulate a matrix multiplication with $2x2$ matrices that contain the relevant information. We will represent the matrices as follows:

$$P_{i,j,\theta} = \begin{bmatrix} c_{ii} & s_{ij} \\ -s_{ji} & c_{jj} \end{bmatrix}, A = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj} \end{bmatrix}, A' = \begin{bmatrix} A'_{ii} & A'_{ij} \\ A'_{ji} & A'_{jj} \end{bmatrix} \tag{3}$$

where the notation A_{ij} denotes element at location (i, j) in matrix A.

Now we will find c, s and θ so that the matrix multiplication nullifies the largest element. In order for P to satisfy eq. 1 we expand the matrix multiplication as follows with $c_{ii} = c_{jj} = c$ and $s_{ji} = s_{ij} = s$:

$$\begin{aligned} \begin{bmatrix} A'_{ii} & A'_{ij} \\ A'_{ji} & A'_{jj} \end{bmatrix} &= \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj} \end{bmatrix} \cdot \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= \begin{bmatrix} c \cdot A_{ii} - s \cdot A_{ji} & c \cdot A_{ij} - s \cdot A_{jj} \\ s \cdot A_{ii} + c \cdot A_{ji} & s \cdot A_{ij} + c \cdot A_{jj} \end{bmatrix} \cdot \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= \begin{bmatrix} c^2 \cdot A_{ii} - cs \cdot A_{ji} - cs \cdot A_{ij} + s^2 \cdot A_{jj} & cs \cdot A_{ii} - s^2 \cdot A_{ji} + c^2 \cdot A_{ij} - cs \cdot A_{jj} \\ cs \cdot A_{ii} + c^2 \cdot A_{ji} - s^2 \cdot A_{ij} - cs \cdot A_{jj} & s^2 \cdot A_{ii} + cs \cdot A_{ji} + cs \cdot A_{ij} + c^2 \cdot A_{jj} \end{bmatrix} \\ &= \begin{bmatrix} c^2 \cdot A_{ii} - 2cs \cdot A_{ij} + s^2 \cdot A_{jj} & (c^2 - s^2) \cdot A_{ij} + cs \cdot (A_{ii} - A_{jj}) \\ (c^2 - s^2) \cdot A_{ij} - cs \cdot (A_{ii} - A_{jj}) & c^2 \cdot A_{jj} + 2cs \cdot A_{ij} + s^2 \cdot A_{ii} \end{bmatrix} \end{aligned} \tag{4}$$

where the last elements are simplified because $A_{ij} = A_{ji}$.

In order to make the non diagonal element in this matrix as 0 we will examine the non diagonal equation as follows:

$$A'_{ij} = (c^2 - s^2) \cdot A_{ij} + cs(A_{ii} - A_{jj}) = 0 \tag{5}$$

Hence it follows that:

$$\frac{c^2 - s^2}{cs} = \frac{A_{jj} - A_{ii}}{A_{ij}} \tag{6}$$

and we can define the rotation angle as follows:

$$\theta = cot(2\phi) = \frac{c^2 - s^2}{2cs} = \frac{A_{jj} - A_{ii}}{2A_{ij}} \tag{7}$$

and by letting $t = s/c$ we can rewrite the equation above as:

$$2cs\theta = c^2 - s^2 <=> t^2 + 2t\theta - 1 = 0 \tag{8}$$

which has the solutions:

$$t = \begin{cases} -\theta + \sqrt{\theta^2 + 1} \\ -(\theta + \sqrt{\theta^2 + 1}) \end{cases} \tag{9}$$

These solutions can be written more succinctly as

$$t = -\theta + \sqrt{\theta^2+1} = \frac{\left(-\theta+\sqrt{\theta^2+1}\right)\left(-\theta+\sqrt{\theta^2+1}\right)}{\theta+\sqrt{\theta^2+1}} = \frac{-\theta^2+\theta^2+1}{\theta+\sqrt{\theta^2+1}} = \frac{1}{\theta+\sqrt{\theta^2+1}} \tag{10}$$

$$t = -\theta - \sqrt{\theta^2+1} = \frac{\left(-\theta-\sqrt{\theta^2+1}\right)\left(\theta-\sqrt{\theta^2+1}\right)}{\theta-\sqrt{\theta^2+1}} = \frac{-\theta^2+\theta^2+1}{\theta-\sqrt{\theta^2+1}} = \frac{-1}{-\theta+\sqrt{\theta^2+1}} \tag{11}$$

We want to rotate the matrix by the angle which corresponds to the smaller root of this equation and generally we can write the smaller root as:

$$t = \frac{sign(\theta)}{|\theta| + \sqrt{1+\theta^2}} \tag{12}$$

and since $t = s/c$ we now have:

$$c = \frac{1}{\sqrt{t^2+1}}, \qquad s = t \cdot c \tag{13}$$

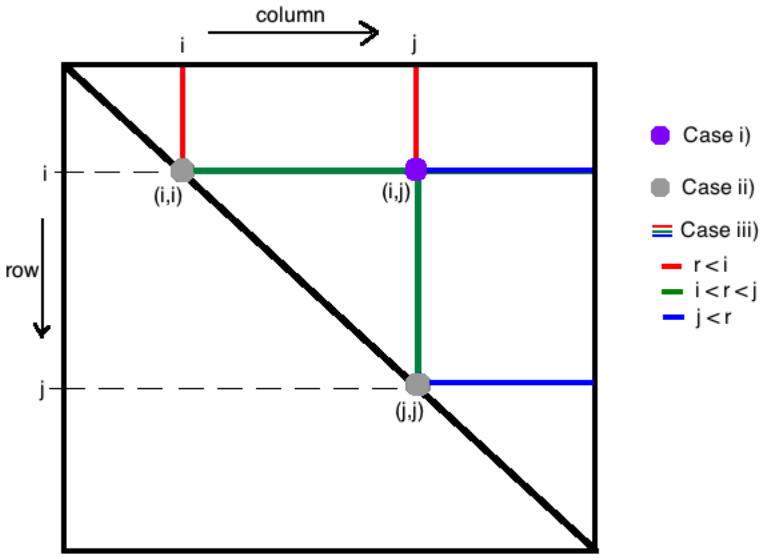
Now we know how to set these variables for the rotations to work and we need to look at three scenarios to update the matrix when a rotation is performed:

i) Set value at location (i,j) as 0

ii) Change diagonal values at locations (i,i) and (j,j)

iii) Change values on rows and columns i and j except (i,i) and (j,j)

The following diagram shows which values are modified during the rotation:



We define a tolerance $tol = 10^{-9}$ and perform rotations until the largest non-diagonal element is less than the tolerance. Since the matrix is symmetric we will perform rotations on the upper triangle of the matrix until it has converged.

For scenario i) we simply set the value of A'_{ij} as 0. However, for scenario ii) we will look at the top left and bottom right elements in eq. 4 to gather equations to set the diagonal elements A'_{ii} and A'_{jj} . We have:

$$A'_{ii} = c^2 \cdot A_{ii} - 2cs \cdot A_{ij} + s^2 \cdot A_{jj} \tag{14}$$

From eq. 5 (because $A'_{ij} = 0$) we can isolate A_{jj} as

$$A_{jj} = A_{ii} - A_{ij}\frac{s^2 - c^2}{cs} \tag{15}$$

and since $c^2 + s^2 = 1$ we simplify eq 14. as:

$$\begin{aligned} A'_{ii} &= c^2 \cdot A_{ii} - 2cs \cdot A_{ij} + s^2 \cdot A_{jj} \\ &= c^2 \cdot A_{ii} - 2cs \cdot A_{ij} + s^2 \left(A_{ii} - A_{ij}\frac{s^2 - c^2}{cs}\right) \\ &= (c^2 + s^2) \cdot A_{ii} - s\left(2c + \frac{s^2 - c^2}{c}\right) A_{ij} \\ &= (c^2 + s^2) \cdot A_{ii} - \frac{s}{c}\left(2c^2 + s^2 - c^2\right) A_{ij} \\ &= A_{ii} - \frac{s}{c}\left(c^2 + s^2\right) A_{ij} \\ &= A_{ii} - t \cdot A_{ij} \end{aligned} \tag{16}$$

Similarly we have:

$$A'_{jj} = A_{jj} + t \cdot A_{ij} \tag{17}$$

For scenario iii) we can look at top of eq 4. and note that if we consider an element A_{rj} when we perform rotation around A_{ij} that only the last two matrices will change the result since the first matrix changes rows i and j and does not have effect on row r. The last matrix changes columns i and j and therefore changes the resulting matrix. Multiplying through these matrices gives us the equations:

$$\begin{cases} A'_{ri} = cA_{ri} - sA_{rj} \\ A'_{rj} = cA_{ri} + sA_{rj} \end{cases} \tag{18}$$

Lets look at A'_{ri} which can be represented as:

$$\begin{aligned} A'_{ri} &= cA_{ri} - sA_{rj} \\ &= \left(1 - \frac{(1-c)(1+c)}{1+c}\right) A_{ri} - sA_{rj} \\ &= \left(1 - \frac{1-c^2}{1+c}\right) A_{ri} - sA_{rj} \\ &= \left(1 - \frac{s^2}{1+c}\right) A_{ri} - sA_{rj} \\ &= A_{ri} - s\left(A_{rj} + \frac{s}{1+c} A_{ri}\right) \\ &= A_{ri} - s\left(A_{rj} + \tau A_{ri}\right) \end{aligned} \tag{19}$$

where

$$\tau = \frac{s}{1+c} \tag{20}$$

which has less roundoff error than eq. 18.

Similarly we have

$$A'_{rj} = A_{rj} + s\left(A_{ri} - \tau A_{rj}\right) \tag{21}$$

To summarise we set values of elements in rows r and l and columns r and l as follows:

i) $A_{ij} = 0$

ii) $\begin{cases} A'_{ii} = A_{ii} - t \cdot A_{ij} \\ A'_{jj} = A_{jj} + t \cdot A_{ij} \end{cases}$

iii) $\begin{cases} A'_{ri} = A_{ri} - s\left(A_{rj} + \tau A_{ri}\right) \\ A'_{rj} = A_{rj} + s\left(A_{ri} - \tau A_{rj}\right) \end{cases}$

where

$$s = t \cdot c, \quad t = \frac{sign(\theta)}{|\theta| + \sqrt{1+\theta^2}}, \quad \tau = \frac{s}{1+c}, \quad \theta = \frac{A_{ii} - A_{jj}}{2A_{ij}}$$