Author: Dr. Shi Guo Email: guoshi1984@hotmail.com

1 Jacobi Method for Solving Eigenvalues

1.1 Intuition

Imagine we have a simple diagonal matrix C

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Finding the eigenvalue and eigenvector is trival. Its eigenvalue is $\lambda_1 = 1$, $\lambda_2 = -1$. The eigenvectors are

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We now consider a rotation matrix that rotates the eigenvectors by 45 degree angle

$$\begin{aligned} u_1 &= Rv_1 \\ &= \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \end{aligned}$$

The eigenequation still holds for u_1 because

$$\begin{aligned} Cv_1 = CR^{-1}Rv_1 &= \lambda v_1 \\ RCR^{-1}Rv_1 &= \lambda Rv_1 \\ RCR^{-1}u_1 &= \lambda u_1 \end{aligned}$$

Let $RCR^{-1} = A$, where A is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Up to now we see we can transform a diagonal matrix to non-diagonal, still symmetric matrix by doing rotation.

1.2 Jacobi Method

The Jacobi method reverse the idea above by rotating a non-diagonal matrix back to a diagonal matrix.

$$\begin{aligned} A &= RCR^{-1} \\ C &= R^{-1}AR \end{aligned}$$

where matrix C is diagonal.

1.3 Eligibility

Since we apply similar transformation by rotation matrix and eventually we can the diagonal matrix which is symmetric, the original matrix has to be symmetrical.

1.4 Algorithm

The Jacobi iteration for a matrix A is

$$A^{(k)} = R^{T}{}^{(k)}_{pq}(\theta)A^{k-1}R^{(k)}_{pq}(\theta)$$

Where

$$R_{pq}(\theta) = \begin{pmatrix} I & 0 & 0 & 0 & 0 \\ 0 & cos(\theta) & 0 & sin(\theta) & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & -sin(\theta) & 0 & cos(\theta) & 0 \\ 0 & 0 & 0 & 0 & I \end{pmatrix}$$

It is an Identity matrix replaced by an rotation matrix on pth and qth columns and rows. The iteration is chosen to reduce the sum of the squares of the off-diagonal elements, which for any square matrix A is

$$||A||_F^2 - \sum_i a_{ii}^2$$

The orthogonal similarity transforms preserve the Frobenius norm

$$||A^{(k)}||_F = ||A^{(k-1)}||_F$$

Because the rotation matrix change only (p,p), (q,q), (p,q), (q,p) positions. We have

$$(a_{pp}^{(k)})^2 + (a_{qq}^{(k)})^2 + 2(a_{pq}^{(k)})^2 = (a_{pp}^{(k-1)})^2 + (a_{qq}^{(k-1)})^2 + 2(a_{pq}^{(k-1)})^2$$

The off-diagonal sum of squares at the kth stage in terms of that at k-1 th stage is

$$\begin{aligned} &||A^{(k)}||_F^2 - \sum_i (a_{ii}^{(k)}) \\ =&||A^{(k)}||_F^2 - \sum_{i \neq p,q} (a_{ii}^{(k)}) - ((a_{pp}^{(k)})^2 + (a_{qq}^{(k)})^2) \\ =&||A^{(k)}||_F^2 - \sum_i (a_{ii}^{(k-1)}) - 2(a_{pq}^{(k-1)})^2 + 2(a_{pq}^{(k)})^2 \end{aligned}$$

In order to minimize this, we need

$$a_{pq}^{(k)} = 0$$
$$a_{pq}^{(k-1)} = max_{i<j}|a_{ij}^{(k-1)}|$$

This implies

$$a_{pq}^{(k-1)}(cos^2\theta - sin^2\theta) + (a_{pp}^{k-1} - a_{qq}^{k-1})cos\theta sin\theta = 0$$

Solve for θ

$$tan(2\theta) = \frac{2a_{pq}^{(k-1)}}{a_{pp}^{k-1} - a_{qq}^{k-1}}$$
$$tan(\theta) = \frac{tan(2\theta)}{1 + \sqrt{1 + tan^2(2\theta)}}$$
$$cos\theta = \frac{1}{\sqrt{1 + tan^2\theta}}$$
$$sin\theta = cos\theta tan\theta$$

We use the above formula to update the matrix(by rotation).

1.5 Eigenvectors

Since

$$A = RCR^{-1}C = R^{-1}AR$$

We know the eigenvectors for diagonal matrix is unit vector e, so

$$Ce = R^{-1}ARe = \lambda e$$

Multiply by R from the left,

$$RPe = \lambda Re$$

so Re is the eigenvector for A.

2 Singular Value Decomposition(SVD)

$$A = U\Sigma V^T$$

2.1 SVD using Jacobi

a. Definition
A matrix $A(m \times n)$ can be decompose into three matrix

$$A = U\Sigma V^T$$

where U is $m \times m$ matrix, Σ is $m \times n$ matrix with diagonal elements only, and V is $n \times n$ matrix. **Jacobi method** SVD using Jacobi is based on the following fact

$$\begin{aligned}
&A = U\Sigma V^T \\
&A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T \\
&A^TAV = V\Sigma^T\Sigma V^T V \\
&A^TAV = V\Sigma^T\Sigma = \Sigma^T\Sigma V
\end{aligned}$$

So V is the eigenvector matrix of A^TA, and $\Sigma^T\Sigma$ is a diagonal matrix whose elements are σ_i^2, where σ_i is the eigenvalue of A. This can be shown as the following.
If $m > n$, then

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix}$$

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

If $m < n$, then

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix}$$

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So σ_i can be calculated by taking the square root of the first r = min(m,n) largest values of σ_i^2.
The U can be obtains in two ways. One is using Jacobi method again.

$$\begin{aligned}
&AA^T = U\Sigma V^T(U\Sigma V^T)^T = U\Sigma V^T V\Sigma U^T = U\Sigma\Sigma^T U^T \\
&AA^TU = U\Sigma\Sigma^T U^T U \\
&AA^TU = \Sigma\Sigma^T U
\end{aligned}$$

So U is the eigenvector matrix of AA^T, but this requires additional Jacobi diagonalization. An alternative way is to consider

$$AV = U\Sigma V^T V = U\Sigma$$

If $m > n$, then

$$U\Sigma = \begin{pmatrix} u_{11}\sigma_1 & u_{12}\sigma_2 \\ u_{21}\sigma_1 & u_{22}\sigma_2 \\ u_{31}\sigma_1 & u_{32}\sigma_2 \end{pmatrix}$$

If $m < n$, then

$$U\Sigma = \begin{pmatrix} u_{11}\sigma_1 & u_{12}\sigma_2 & 0 \\ u_{21}\sigma_1 & u_{22}\sigma_2 & 0 \end{pmatrix}$$

So if we take $U\Sigma$ matrix, divided by the ith column with σ_i, we can obtain U matrix.