Lecture 1

SVD and PCA

SVD

- A form of matrix decomposition (factorization into a product of matrices)
- Other forms of matrix decomposition (Eigendecomposition, LU factorization/reduction--should be familiar from Math 54)
- In singular value decomposition, the matrix can be written as a sum of rank-1 matrices
- A = σ1u1vT1 + σ2u2vT2 + . . . + σnunvT (format later)
- Each one of these matrices is a mode.
- σ values: **the singular values are square roots of eigenvalues from AA^T or A^TA.**

Numerical Example

Example: Find the SVD of A, $U\Sigma V^T$, where $A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$.

First we compute the singular values σ_i by finding the eigenvalues of AA^T.

$$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}.$$

The characteristic polynomial is $det(AA^T - \lambda I) = \lambda^2 - 34\lambda + 225 = (\lambda - 25)(\lambda - 9)$, so the singular values are $\sigma_1 = \sqrt{25} = 5$ and $\sigma_2 = \sqrt{9} = 3$.

Now we find the right singular vectors (the columns of V) by finding an orthonormal set of eigenvectors of A^TA. It is also possible to proceed by finding the left singular vectors (columns of U) instead. The eigenvalues of A^TA are 25, 9, and 0, and since A^TA is symmetric we know that the eigenvectors will be orthogonal.

For $\lambda = 25$, we have

$$A^T A - 25I = \begin{pmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{pmatrix}$$

which row-reduces to $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. A unit-length vector in the kernel of that matrix is $v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$.

For $\lambda = 9$ we have $A^T A - 9I = \begin{pmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{pmatrix}$ which row-reduces to $\begin{pmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{pmatrix}$.

A unit-length vector in the kernel is $v_2 = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix}$.

For the last eigenvector, we could compute the kernel of $A^T A$ or find a unit vector perpendicular to v_1 and v_2. To be perpendicular to $v_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ we need $-a = b$. Then the condition that $v_2^T v_3 = 0$ becomes $2a/\sqrt{18} + 4c/\sqrt{18} = 0$ or $-a = 2c$. So $v_3 = \begin{pmatrix} a \\ -a \\ -a/2 \end{pmatrix}$ and for it to be unit-length we need $a = 2/3$ so $v_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$.

So at this point we know that

$$A = U\Sigma V^T = U \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}.$$

Finally, we can compute U by the formula $\sigma u_i = Av_i$, or $u_i = \frac{1}{\sigma} Av_i$. This gives $U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$. So in its full glory the SVD is:

$$A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}.$$

Applications

Some applications of the SVD:

- Computing the pseudoinverse of a matrix
- matrix approximation
- determining the rank, range and null space of a matrix.
- separable models.
- nearest orthogonal matrix.
- Kabsch algorithm (calculating the optimal rotation matrix that minimizes the root mean squared deviation between two paired sets of points)

PCA

- dimensionality-reduction method: for analysis, it is often necessary to reduce the dimensionality of large data sets
- transforming a large set of variables into a smaller one that still contains most of the information in the large set.
- small loss of accuracy, but the aim is to achieve a simpler dataset by preserving as much information as possible.
- Steps: standardization, covariance matrix computation, and identify principal components by computing the eigenvalues of the covariance matrix.
- For standardization: subtract the mean and divide by the standard deviation for each value of each variable.

Standardization and computing the covariance matrix

$$z = \frac{value - mean}{standard\ deviation}$$

$$\begin{bmatrix} Cov(x,x) & Cov(x,y) & Cov(x,z) \\ Cov(y,x) & Cov(y,y) & Cov(y,z) \\ Cov(z,x) & Cov(z,y) & Cov(z,z) \end{bmatrix}$$

Applications

Some applications of PCA:

- Computing the pseudoinverse of a matrix
- dimensionality reduction
- multivariate analysis. E
- data compression
- image processing
- Visualization
- exploratory data analysis,
- pattern recognition and time series prediction.