Singular Value Decomposition of a Rectangular Matrix

Singular Value Decomposition

The singular value decomposition and the pursuing eigenanalysis is generally applied to a square matrix. In this presentation, the eigenanalysis is used to develop a similar decomposition, called the singular value decomposition, for a rectangular matrix. The singular value decomposition is then used to give the principal component analysis.

Let $\mathbf{X}$ be an $n \times p$ matrix with $n > p$. Then $\mathbf{X}'\mathbf{X}$ is a square symmetric matrix of order $p \times p$. $\mathbf{X}'\mathbf{X}$ can be expressed in terms of its eigenvalues $\mathbf{L}$ and eigenvectors $\mathbf{Z}$ as

$$\mathbf{X}'\mathbf{X} = \mathbf{Z}\mathbf{L}\mathbf{Z}'$$

Similarly, $\mathbf{X}\mathbf{X}'$ is a square symmetric matrix but of order $n \times n$. The rank of $\mathbf{X}\mathbf{X}'$ will be at most p so there will be at most p nonzero eigenvalues; they are in fact the same p eigenvalues obtained from $\mathbf{X}'\mathbf{X}$ In addition, $\mathbf{X}'\mathbf{X}$ will have n - p eigenvalues that are zero. These n - p eigenvalues and their vectors are dropped in the following discussion. Denote with $\mathbf{U}$ the matrix of eigenvectors of $\mathbf{X}\mathbf{X}'$ that correspond to the p eigenvalues common to $\mathbf{X}'\mathbf{X}$. Each eigenvector $\mathbf{u}$; will be of order $n \times 1$. Then,

$$\mathbf{X}\mathbf{X}' = \mathbf{U}\mathbf{L}\mathbf{U}'$$

The rectangular matrix $\mathbf{X}$ can be written as

$$\mathbf{X} = \mathbf{U}\mathbf{L}^{\frac{1}{2}}\mathbf{Z}' \tag{1}$$

where $\mathbf{L}^{\frac{1}{2}}$ is the diagonal matrix of the positive square roots of the p eigenvalues of $\mathbf{X}'\mathbf{X}$ Thus, $\mathbf{L}^{\frac{1}{2}}\mathbf{L}^{\frac{1}{2}} = \mathbf{L}$. Equation (1) is the singular value decomposition of the rectangular matrix $\mathbf{X}$ The elements of $\mathbf{L}^{\frac{1}{2}}$, $\sqrt{\lambda_i}$, are called the singular values and the column vectors in $\mathbf{U}$ and $\mathbf{Z}$ are the left and right singular vectors, respectively.

Since $\mathbf{L}^{\frac{1}{2}}$ is a diagonal matrix, the singular value decomposition expresses $\mathbf{X}$ as a sum of p rank-1 matrices,

$$\mathbf{X} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{u}_i z_i'$$

where summation is over $i = 1,\ldots,p$. Furthermore, if the eigenvalues have been ranked from largest to smallest, the first of these matrices is the "best" rank-1 approximation to $\mathbf{X}$, the sum of the first two matrices is the "best" rank-2 approximation of $\mathbf{X}$, and so forth. These are "best" approximations in the least squares sense; that is, no other matrix (of the same rank) will give a better agreement with the original matrix $\mathbf{X}$ as measured by the sum of squared differences between the corresponding elements of $\mathbf{X}$ and the approximating matrix (Householder and Young, 1938). The goodness of fit of the approximation in each case is given by the ratio of the sum of the eigenvalues (squares of the singular values) used in the approximation to the sum of all eigenvalues. Thus, the rank-1 approximation has a goodness of fit of $\frac{\lambda_1}{\sum \lambda_i}$, the rank-2 approximation has a goodness of fit of $\frac{\lambda_1+\lambda_2}{\sum \lambda_i}$, and so forth.

Singular value decomposition is illustrated using data on average minimum daily temperature, X_1, average maximum daily temperature, X_2, total rainfall, X_3, and total growing degree days, X_4, for six locations. The data were reported by Saeed and Francis (1984) to relate environmental conditions to cultivar-by- environment interactions in sorghum and are used with their kind permission. Below is the original data.

Average temperature

	Minimum		*Maximum*		*Total rain*		*Total GDD*	
Location	1978	1979	1978	1979	1978	1979	1978	1979
			C			*-cm-*		
Concord	11.1	10.6	25.2	24.1	33.0	63.4	1576	1438
Mead	12.6	11.9	26.0	27.7	48.2	42.8	1767	1831
Tyson	9.3		27.3		27.3		1580	
Scottsbluff	9.8	10.4	26.7	26.5	34.3	28.7	1578	1614
Sidney	5.8	6.7	25.7	25.4	27.3	28.2	1222	1232
Garden City	12.1	11.8	28.3	28.4	20.4	64.2	1926	1876

Each variable has been centered to have zero mean, and standardized to have unit sum of squares. (The centering and standardization are not necessary for a singular value decomposition. The centering removes the mean effect of each variable so that the dispersion about the mean is being analyzed. The standardization puts all variables on an equal basis and is desirable in most cases, particularly when the variables have different units of measure.) The $\mathbf{X}$ matrix is

$$\mathbf{X} = \begin{bmatrix} .178146 & -.523245 & .059117 & -.060996 \\ .449895 & -.209298 & .777976 & .301186 \\ -.147952 & .300866 & -.210455 & -.053411 \\ -.057369 & .065406 & .120598 & -.057203 \\ -.782003 & -.327028 & -.210455 & -.732264 \\ .359312 & .693299 & -.536780 & .602687 \end{bmatrix}$$

The singular value decomposition of $\mathbf{X}$ into $\mathbf{U}\mathbf{L}^{\frac{1}{2}}\mathbf{Z}'$ gives

$$\mathbf{U} = \begin{bmatrix} -.113995 & .308905 & -.810678 & .260088 \\ .251977 & .707512 & .339701 & -.319261 \\ .007580 & -.303203 & .277432 & .568364 \\ -.028067 & .027767 & .326626 & .357124 \\ -.735417 & -.234888 & .065551 & -.481125 \\ .617923 & -.506093 & -.198632 & -.385189 \end{bmatrix}$$

$$\mathbf{L}^{\frac{1}{2}*} = \begin{bmatrix} 1.496896 & 0 & 0 & 0 \\ 0 & 1.244892 & 0 & 0 \\ 0 & 0 & .454086 & 0 \\ 0 & 0 & 0 & .057893 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} .595025 & .336131 & -.383204 & .621382 \\ .451776 & -.540753 & .657957 & .265663 \\ .004942 & .768694 & .639051 & -.026450 \\ .664695 & .060922 & -.108909 & -.736619 \end{bmatrix}.$$

The columns of $\mathbf{U}$ and $\mathbf{Z}$ are the left and right singular vectors, respectively. The first column of $\mathbf{U}$, $\mathbf{u}_1$, the first column of $\mathbf{Z}$, $\mathbf{z}_1$, and the first singular value, $\sqrt{\lambda_1}$= 1.496896, or and eigen value of λ_1 = give the best rank-1 approximation of $\mathbf{X}$

$$\begin{aligned} \mathbf{X}_1 &= \sqrt{\lambda_1}\mathbf{u}_1 z_1' \\ &= (1.4969) \begin{pmatrix} -.113995 \\ .251977 \\ .007580 \\ -.028067 \\ -.735417 \\ .617923 \end{pmatrix} \begin{pmatrix} .595025 & .451776 & .004942 & .664695 \end{pmatrix} \\ &= \begin{bmatrix} -.101535 & -.077091 & -.000843 & -.113423 \\ .224434 & .170403 & .001864 & .250712 \\ .006752 & .005126 & .000056 & .007542 \\ -.024999 & -.018981 & -.000208 & -.027927 \\ -.655029 & -.497335 & -.005440 & -.731725 \\ .550378 & .417877 & .004571 & .614820 \end{bmatrix}. \end{aligned}$$

The goodness of fit of $\mathbf{X}_1$ to $\mathbf{X}$ is measured by

$$\begin{aligned} \frac{\lambda_1}{\sum_{i=1}^{4}\lambda_i} &= \frac{1.4969^2}{4} \\ &= 0.56 \end{aligned}$$

or the sum of squares of the differences between the elements of $\mathbf{X}$ and $\mathbf{X}_1$, the lack of fit, is 44% of the total sum of squares of the elements in $\mathbf{X}$. This is not a very good approximation. The rank-2 approximation to $\mathbf{X}$ is obtained by adding to $\mathbf{X}_1$ the matrix $\mathbf{X}_2$.

$$\begin{aligned} \mathbf{X}_2 &= \sqrt{\lambda_2}\mathbf{u}_2 z_2' \\ &= (1.2449) \begin{pmatrix} .308905 \\ .707512 \\ -.303203 \\ .027767 \\ -.234888 \\ -.506093 \end{pmatrix} \begin{pmatrix} .336131 & -.540753 & .768694 & .060922 \end{pmatrix} \end{aligned}$$

$$= \begin{bmatrix} 0.1293 & -0.2079 & 0.2956 & 0.0234 \\ 0.2961 & -0.4763 & 0.6771 & 0.0537 \\ -0.1269 & 0.2041 & -0.2901 & -0.0230 \\ 0.0116 & -0.0187 & 0.0266 & 0.0021 \\ -0.0983 & 0.1581 & -0.2248 & -0.0178 \\ -0.2118 & 0.3407 & -0.4843 & -0.0384 \end{bmatrix}$$

Which gives

$$\mathbf{X}_1 + \mathbf{X}_2 = \begin{bmatrix} 0.0277 & -0.2850 & 0.2948 & -0.0900 \\ 0.5205 & -0.3059 & 0.6789 & 0.3044 \\ -0.1201 & 0.2092 & -0.2901 & -0.0155 \\ -0.0134 & -0.0377 & 0.0264 & -0.0258 \\ -0.7533 & -0.3392 & -0.2302 & -0.7495 \\ 0.3386 & 0.7586 & -0.4797 & 0.5764 \end{bmatrix}$$

which has goodness of fit

$$\begin{aligned} \frac{\lambda_1 + \lambda_2}{\sum_{i=1}^{4} \lambda_i} &= \frac{1.4969^2 + 1.2448^2}{4} \\ &= 0.95 \end{aligned}$$

In terms of approximating $\mathbf{X}$ with the rank-2 matrix $\mathbf{X}_1 + \mathbf{X}_2$, the goodness of fit of .95 means that the sum of squares of the discrepancies between $\mathbf{X}$ and ($\mathbf{X}_1 + \mathbf{X}_2$,) is 5% of the total sum of squares of all elements in $\mathbf{X}$. The sum of squares of all elements in $\mathbf{X}$ is $\sum_{i=1}^{4} \lambda_i$, the sum of squares of all elements in ($\mathbf{X}_1 + \mathbf{X}_2$,) is $(\lambda_1 + \lambda_2)$, and the sum of squares of all elements in [$\mathbf{X}$-($\mathbf{X}_1 + \mathbf{X}_2$,)] is $(\lambda_3 + \lambda_4)$. In terms of the geometry of the data vectors, the goodness of fit of .95 means that 95% of the dispersion of the "cloud" of points in the original four-dimensional space is, in reality, contained in two dimensions, or the points in four-dimensional space very nearly fall on a plane. Only 5% of the dispersion is lost if the third and fourth dimensions are ignored. Using all four singular values and their singular vectors gives the complete decomposition of $\mathbf{X}$ into four orthogonal rank-1 matrices. The sum of the four matrices equals $\mathbf{X}$ within the limits of rounding error. The analysis has shown, by the relatively small size of the third and fourth singular values, that the last two dimensions contain little of the dispersion and can safely be ignored in interpretation of the data.

```
X =

    0.1781   -0.5232    0.0591   -0.0610
    0.4499   -0.2093    0.7780    0.3012
   -0.1480    0.3009   -0.2105   -0.0534
   -0.0574    0.0654    0.1206   -0.0572
   -0.7820   -0.3270   -0.2105   -0.7323
    0.3593    0.6933   -0.5368    0.6027

svd(X)

ans =
    1.4969
    1.2449
    0.4541
    0.0579

[U,L,Z]=svd(X)

U =
   -0.1140   -0.3089    0.8107   -0.2602    0.4071    0.0311
    0.2520   -0.7075   -0.3397    0.3192    0.4284   -0.1873
    0.0076    0.3032   -0.2774   -0.5684    0.4628   -0.5420
   -0.0281   -0.0278   -0.3266   -0.3572    0.3314    0.8089
   -0.7354    0.2349   -0.0655    0.4811    0.4102    0.0004
    0.6179    0.5061    0.1986    0.3851    0.3980    0.1260


L =
    1.4969         0         0         0
         0    1.2449         0         0
         0         0    0.4541         0
         0         0         0    0.0579
         0         0         0         0
         0         0         0         0


Z =
    0.5950   -0.3361    0.3832   -0.6214
    0.4518    0.5407   -0.6580   -0.2657
    0.0049   -0.7687   -0.6390    0.0264
    0.6647   -0.0609    0.1089    0.7366

X1=L(1:1,1:1)*U(1:6,1:1)*Z(1:4,1:1)'

X1 =
```

```
    -0.1015   -0.0771   -0.0008   -0.1134
     0.2244    0.1704    0.0019    0.2507
     0.0068    0.0051    0.0001    0.0075
    -0.0250   -0.0190   -0.0002   -0.0279
    -0.6550   -0.4973   -0.0054   -0.7317
     0.5504    0.4179    0.0046    0.6148

X2=L(2:2,2:2)*U(1:6,2:2)*Z(1:4,2:2)'

X2 =
     0.1293   -0.2079    0.2956    0.0234
     0.2961   -0.4763    0.6771    0.0537
    -0.1269    0.2041   -0.2901   -0.0230
     0.0116   -0.0187    0.0266    0.0021
    -0.0983    0.1581   -0.2248   -0.0178
    -0.2118    0.3407   -0.4843   -0.0384

X3=L(3:3,3:3)*U(1:6,3:3)*Z(1:4,3:3)'

X3 =
     0.1411   -0.2422   -0.2352    0.0401
    -0.0591    0.1015    0.0986   -0.0168
    -0.0483    0.0829    0.0805   -0.0137
    -0.0568    0.0976    0.0948   -0.0162
    -0.0114    0.0196    0.0190   -0.0032
     0.0346   -0.0593   -0.0576    0.0098

X4=L(4:4,4:4)*U(1:6,4:4)*Z(1:4,4:4)'

X4 =
     0.0094    0.0040   -0.0004   -0.0111
    -0.0115   -0.0049    0.0005    0.0136
     0.0204    0.0087   -0.0009   -0.0242
     0.0128    0.0055   -0.0005   -0.0152
    -0.0173   -0.0074    0.0007    0.0205
    -0.0139   -0.0059    0.0006    0.0164

X1+X2

ans =
     0.0277   -0.2850    0.2948   -0.0900
     0.5205   -0.3059    0.6789    0.3044
    -0.1201    0.2092   -0.2901   -0.0155
    -0.0134   -0.0377    0.0264   -0.0258
    -0.7533   -0.3392   -0.2302   -0.7495
     0.3386    0.7586   -0.4797    0.5764
```

```
X1+X2+X3

ans =
     0.1688   -0.5272    0.0595   -0.0499
     0.4614   -0.2044    0.7775    0.2876
    -0.1684    0.2921   -0.2096   -0.0292
    -0.0702    0.0599    0.1211   -0.0420
    -0.7647   -0.3196   -0.2112   -0.7528
     0.3732    0.6992   -0.5374    0.5863

X1+X2+X3+X4   (Which is the original X-matrix )

ans =
     0.1781   -0.5232    0.0591   -0.0610
     0.4499   -0.2093    0.7780    0.3012
    -0.1480    0.3009   -0.2105   -0.0534
    -0.0574    0.0654    0.1206   -0.0572
    -0.7820   -0.3270   -0.2105   -0.7323
     0.3593    0.6933   -0.5368    0.6027

L =
     1.4969         0         0         0
          0    1.2449         0         0
          0         0    0.4541         0
          0         0         0    0.0579
          0         0         0         0
          0         0         0         0

L'*L (Eigenvalues)

ans =
     2.2407         0         0         0
          0    1.5498         0         0
          0         0    0.2062         0
          0         0         0    0.0034
```