The Singular Value Decomposition of $\boldsymbol{X}$

Stephen Vardeman

Analytics Iowa LLC

ISU Statistics and IMSE

Singular vectors and values

If the $N \times p$ matrix $\mathbf{X}$ has rank r then it has a so-called singular value decomposition (SVD) as

$$\underset{N\times p}{\mathbf{X}} = \underset{N\times r}{\mathbf{U}}\ \underset{r\times r}{\mathbf{D}}\ \underset{r\times p}{\mathbf{V}'}$$

where $\mathbf{U}$ has orthonormal columns (left singular vectors) spanning $C(\mathbf{X})$, $\mathbf{V}$ has orthonormal columns (right singular vectors) spanning $C(\mathbf{X}')$ (the row space of $\mathbf{X}$), and $\mathbf{D} = \mathbf{diag}(d_1, d_2, \ldots, d_r)$ for

$$d_1 \geq d_2 \geq \cdots \geq d_r > 0$$

The d_j are the "singular values" of $\mathbf{X}$. (For a real non-negative definite square matrix–a covariance matrix–the SVD is the eigen decomposition, $\mathbf{U} = \mathbf{V}$, columns of these matrices are unit eigenvectors, and the SVD singular values are the corresponding eigenvalues.)

Low rank approximations to *X* matrices

If $\mathbf{U}_l$ and $\mathbf{V}_l$ are matrices consisting of the first $l \leq r$ columns of $\mathbf{U}$ and $\mathbf{V}$ respectively, then

$$\mathbf{X}^{*l} = \mathbf{U}_l \, \mathbf{diag}\left(d_1, d_2, \ldots, d_l\right) \mathbf{V}_l'$$

is the best (in the sense of squared distance from $\mathbf{X}$ in $\Re^{Np}$) $rank = l$ approximation to $\mathbf{X}$. (Application of this kind of argument to covariance matrices provides low-rank approximations to complicated covariance matrices.)

SVD and OLS

Since the columns of $\mathbf{U}$ are an orthonormal basis for $C(\mathbf{X})$, the projection of an output vector $\mathbf{Y}$ onto $C(\mathbf{X})$ is

$$\widehat{\mathbf{Y}}^{\text{OLS}} = \sum_{j=1}^{r} \langle \mathbf{Y}, \mathbf{u}_j \rangle \mathbf{u}_j = \mathbf{U}\mathbf{U}'\mathbf{Y}$$

In the full rank ($rank = p$) $\mathbf{X}$ case, this is completely parallel to $\widehat{\mathbf{Y}}^{\text{OLS}} = \sum_{j=1}^{r} \langle \mathbf{Y}, \mathbf{q}_j \rangle \mathbf{q}_j = \mathbf{Q}\mathbf{Q}'\mathbf{Y}$ (for $\mathbf{Q}$ from the QR decomposition) and is a consequence of the fact that the columns of *both* $\mathbf{U}$ and $\mathbf{Q}$ form orthonormal bases for $C(\mathbf{X})$. In general, the two bases are not the same.

SVD and *X'X*

Now using the SVD of a full rank (rank p) $\mathbf{X}$,

$$\mathbf{X}'\mathbf{X} = \mathbf{V}\mathbf{D}'\mathbf{U}'\mathbf{U}\mathbf{D}\mathbf{V}' = \mathbf{V}\mathbf{D}^2\mathbf{V}'$$

which is the eigen (or spectral) decomposition of the symmetric and positive definite $\mathbf{X}'\mathbf{X}$. Eigenvalues are the squares of the SVD singular values for $\mathbf{X}$.

Geometry and the SVD of *X*

The vector

$$\mathbf{z}_1 \equiv \mathbf{X}\mathbf{v}_1$$

is the product $\mathbf{X}\mathbf{w}$ with the largest squared length in $\Re^N$ subject to the constraint that $\|\mathbf{w}\| = 1$. Further,

$$\mathbf{z}_1 = \mathbf{X}\mathbf{v}_1 = \mathbf{U}\mathbf{D}\mathbf{V}'\mathbf{v}_1 = \mathbf{U}\mathbf{D}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = d_1\mathbf{u}_1$$

and we see that this squared length is d_1^2 and that this vector points in the direction of $\mathbf{u}_1$.

Geometry and the SVD of *X* cont.

In general,

$$\mathbf{z}_j = \mathbf{X}\mathbf{v}_j = \begin{pmatrix} \langle \mathbf{x}_1, \mathbf{v}_j \rangle \\ \vdots \\ \langle \mathbf{x}_N, \mathbf{v}_j \rangle \end{pmatrix} = d_j \mathbf{u}_j$$

is the vector $\mathbf{X}\mathbf{w}$ with the largest squared length in $\Re^N$ subject to the constraints that $\|\mathbf{w}\| = 1$ and $\langle \mathbf{w}, \mathbf{z}_l \rangle = 0$ for all $l < j$. This squared length is d_j^2 and this vector points in the direction of $\mathbf{u}_j$.