WIKIPEDIA

奇异值分解

维基百科，自由的百科全书

奇异值分解（singular value decomposition）是线性代数中一种重要的矩阵分解，在信号处理、统计学等领域有重要应用。奇异值分解在某些方面与对称矩阵或厄米矩阵基于特征向量的对角化类似。然而这两种矩阵分解尽管有其相关性，但还是有明显的不同。对称阵特征向量分解的基础是谱分析，而奇异值分解则是谱分析理论在任意矩阵上的推广。

目录

理论描述

假设M是一个$m \times n$阶矩阵，其中的元素全部属于域K，也就是实数域或复数域。如此则存在一个分解使得

$$M = U\Sigma V^*,$$

其中U是$m \times m$阶酉矩阵；Σ是$m \times n$阶非负实数对角矩阵；而V^*，即V的共轭转置，是$n \times n$阶酉矩阵。这样的分解就称作M的**奇异值分解**。Σ对角线上的元素$\Sigma_{i,i}$即为M的**奇异值**。

常见的做法是将奇异值由大而小排列。如此Σ便能由M唯一确定了。（虽然U和V仍然不能确定。）

直观的解释

在矩阵M的奇异值分解中

$$M = U\Sigma V^*,$$

- V的列（rows）组成一套对M 的正交"输入"或"分析"的基向量。这些向量是$M^* M$的特征向量。
- U的列（rows）组成一套对M 的正交"输出"的基向量。这些向量是MM^* 的特征向量。
- Σ对角线上的元素是奇异值，可视为是在输入与输出间进行的标量的"膨胀控制"。这些是MM^* 及$M^* M$的特征值的非负平方根，并与U和 V的行向量相对应。

奇异值和奇异向量，以及他们与奇异值分解的关系

一个非负实数σ是M的一个奇异值仅当存在K^m的单位向量u和K^n的单位向量v如下：

$$Mv = \sigma u \text{ and } M^* u = \sigma v.$$

其中向量u和v分别为σ的左奇异向量和右奇异向量。

对于任意的奇异值分解

$$M = U\Sigma V^*$$

矩阵Σ的对角线上的元素等于M的奇异值. U和V的列分别是奇异值中的左、右奇异向量。因此，上述定理表明：

- 一个m×n的矩阵至多有p = min(m,n)个不同的奇异值；
- 总能在K^m中找到由M的左奇异向量组成的一组正交基U，；
- 总能在K^n找到由M的右奇异向量组成的一组正交基V，。

如果对于一个奇异值，可以找到两组线性无关的左（右）奇异向量，则该奇异值称为简并的（或退化的）。

非退化的奇异值在最多相差一个相位因子$\exp(i\phi)$（若讨论限定在实数域内，则最多相差一个正负号）的意义下具有唯一的左、右奇异向量。因此，如果M的所有奇异值都是非退化且非零，则除去一个可以同时乘在U, V上的任意的相位因子外，M的奇异值分解唯一。

根据定义，退化的奇异值具有不唯一的奇异向量。因为，如果u_1和u_2为奇异值σ的两个左奇异向量，则它们的任意归一化线性组合也是奇异值σ一个左奇异向量，右奇异向量也具有类似的性质。因此，如果M具有退化的奇异值，则它的奇异值分解是不唯一的。

例子

观察一个4×5的矩阵

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \end{bmatrix}.$$

M矩阵的奇异值分解如下$U\Sigma V^*$

$$U = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, V^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ \sqrt{0.8} & 0 & 0 & 0 & -\sqrt{0.2} \end{bmatrix}.$$

注意矩阵Σ的所有非对角元为0。矩阵U和V^*都是酉矩阵，它们乘上各自的共轭转置都得到单位矩阵。如下所示。在这个例子中，由于U和V^*都是实矩阵，故它们都是正交矩阵。

$$UU^* = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv I_4$$

$$VV^* = \begin{bmatrix} 0 & 0 & \sqrt{0.2} & 0 & \sqrt{0.8} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{0.8} & 0 & -\sqrt{0.2} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ \sqrt{0.8} & 0 & 0 & 0 & -\sqrt{0.2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \equiv I_5.$$

由于$\mathbf{\Sigma}$有一个对角元是零，故这个奇异值分解值不是唯一的。例如，选择V使得

$$V^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ \sqrt{0.4} & 0 & 0 & \sqrt{0.5} & -\sqrt{0.1} \\ -\sqrt{0.4} & 0 & 0 & \sqrt{0.5} & \sqrt{0.1} \end{bmatrix}$$

能得到M的另一个奇异值分解。

与特征值分解的联系

奇异值分解能够用于任意$m \times n$矩阵，而特征分解只能适用于特定类型的方阵，故奇异值分解的适用范围更广。不过，这两个分解之间是有关联的。给定一个M的奇异值分解，根据上面的论述，两者的关系式如下：

$$M^*M = V\Sigma^*U^*\,U\Sigma V^* = V(\Sigma^*\Sigma)V^*$$

$$MM^* = U\Sigma V^*\,V\Sigma^*U^* = U(\Sigma\Sigma^*)U^*$$

关系式的右边描述了关系式左边的特征值分解。于是：

- V的列向量（右奇异向量）是M^*M的特征向量。
- U的列向量（左奇异向量）是MM^*的特征向量。
- Σ的非零对角元（非零奇异值）是M^*M或者MM^*的非零特征值的平方根。

特殊情况下，当M是一个正规矩阵（因而必须是方阵）根据谱定理，M可以被一组特征向量酉对角化，所以它可以表为：

$$M = UDU^*$$

其中U为一个酉矩阵，D为一个对角阵。如果M是半正定的，$M = UDU^*$的分解也是一个奇异值分解。

然而，一般矩阵的特征分解跟奇异值分解不同。特征分解如下：

$$M = UDU^{-1}$$

其中U是不需要是酉的，D也不需要是半正定的。而奇异值分解如下：

$$M = U\Sigma V^*$$

其中$\mathbf{\Sigma}$是对角半正定矩阵，U和V是酉矩阵，两者除了通过矩阵M没有必然的联系。

几何意义

因为U和V都是酉的，我们知道U的列向量$u_1,...,u_m$组成了K^m空间的一组标准正交基。同样，V的列向量$v_1,...,v_n$也组成了K^n空间的一组标准正交基（根据向量空间的标准点积法则）。

矩阵M代表从K^n到K^m的一个线性映射$\mathcal{T}$：$x \to Mx$。通过这些标准正交基，这个变换可以用很简单的方式进行描述：$\mathcal{T}(v_i) = \sigma_i u_i, i = 1, \ldots, \min(m, n)$，其中$\sigma_i$是$\mathbf{\Sigma}$中的第$i$个对角元。当$i > \min(m, n)$时，$\mathcal{T}(v_i) = 0$。

这样，SVD分解的几何意义就可以做如下的归纳：对于每一个线性映射$\mathcal{T}: K^n \to K^m$，$\mathcal{T}$的奇异值分解在原空间与像空间中分别找到一组标准正交基，使得$\mathcal{T}$把K^n的第i个基向量映射为K^m的第i个基向量的非负倍数，并将K^n中余下的基向量映射为零向量。换句话说，线性变换$\mathcal{T}$在这两组选定的基上的矩阵表示为所有对角元均为非负数的对角矩阵。

应用

求广义逆阵（伪逆）

奇异值分解可以被用来计算矩阵的广义逆阵（伪逆）。若矩阵M的奇异值分解为$M = U\Sigma V^*$，那么M的伪逆为

$$M^+ = V\Sigma^+ U^*,$$

其中Σ^+是Σ的伪逆，是将Σ主对角线上每个非零元素都求倒数之后再转置得到的。求伪逆通常可以用来求解最小二乘法问题。

列空间、零空间和秩

奇异值分解的另一个应用是给出矩阵的列空间、零空间和秩的表示。对角矩阵Σ的非零对角元素的个数对应于矩阵M的秩。与零奇异值对应的右奇异向量生成矩阵M的零空间，与非零奇异值对应的左奇异向量则生成矩阵M的列空间。在线性代数数值计算中奇异值分解一般用于确定矩阵的有效秩，这是因为，由于舍入误差，秩亏矩阵的零奇异值可能会表现为很接近零的非零值。

矩阵近似值

奇异值分解在统计中的主要应用为主成分分析（PCA）。数据集的特征值（在SVD中用奇异值表征）按照重要性排列，降维的过程就是舍弃不重要的特征向量的过程，而剩下的特征向量张成空间为降维后的空间。

几种编程语言中计算SVD的函式范例

- matlab：

  [b c d]=svd(x)

- OpenCV:

  void cvSVD(CvArr* A, CvArr* W, CvArr* U=NULL, CvArr* V=NULL, int flags=0)

- Python（使用SciPy (https://www.scipy.org/)库）

  U,s,Vh = scipy.linalg.svd(A)

- R:

  S=svd(x)

历史

参见

外部链接

- LAPACK users manual (http://www.netlib.org/lapack/lug/node53.html) gives details of subroutines to calculate the SVD (see also [1] (http://www.netlib.org/lapack/lug/node32.html)).
- Applications of SVD (https://web.archive.org/web/20050820034632/http://www2.imm.dtu.dk/~pch/Projekter/tsvd.html) on PC Hansen's web site.
- Introduction to the Singular Value Decomposition (https://web.archive.org/web/20060421024156/http://www.uwlax.edu/faculty/will/svd/) by Todd Will of the University of Wisconsin--La Crosse.
- Los Alamos group's book chapter (http://public.lanl.gov/mewall/kluwer2002.html) has helpful gene data analysis examples.
- MIT Lecture (http://ocw.mit.edu/OcwWeb/Mathematics/18-06Spring-2005/VideoLectures/index.htm) series by Gilbert Strang. See Lecture #29 on the SVD.
- Java SVD (https://web.archive.org/web/20060313144512/http://www.idiom.com/~zilla/Computer/Javanumeric/index.html) library routine.
- Java applet (https://web.archive.org/web/20051222074050/http://klebanov.homeip.net/~pavel/fb/java/la_applets/SVD/index.html) demonstrating the SVD.
- Java script (http://users.pandora.be/paul.larmuseau/SVD.htm) demonstrating the SVD more extensively, paste your data from a spreadsheet.
- Chapter from "Numerical Recipes in C" (http://www.library.cornell.edu/nr/bookcpdf/c2-6.pdf) gives more information about implementation and applications of SVD.
- Online Matrix Calculator (http://www.bluebit.gr/matrix-calculator/) Performs singular value decomposition of matrices.

参考文献

- Demmel, J. and Kahan, W. (1990). Computing Small Singular Values of Bidiagonal Matrices With Guaranteed High Relative Accuracy. *SIAM J. Sci. Statist. Comput.*, **11** (5), 873-912.
- Golub, G. H. and Van Loan, C. F. (1996). "Matrix Computations". 3rd ed., Johns Hopkins University Press, Baltimore. ISBN 0-8018-5414-8.
- Halldor, Bjornsson and Venegas, Silvia A. (1997). "A manual for EOF and SVD analyses of climate data" (http://www.vedur.is/~halldor/TEXT/eofsvd.html). McGill University, CCGCR Report No. 97-1, Montréal, Québec, 52pp.
- Hansen, P. C. (1987). The truncated SVD as a method for regularization. *BIT*, **27**, 534-553.
- Horn, Roger A. and Johnson, Charles R (1985). "Matrix Analysis". Section 7.3. Cambridge University Press. ISBN 0-521-38632-2.
- Horn, Roger A. and Johnson, Charles R (1991). Topics in Matrix Analysis, Chapter 3. Cambridge University Press. ISBN 0-521-46713-6.
- Strang G (1998). "Introduction to Linear Algebra". Section 6.7. 3rd ed., Wellesley-Cambridge Press. ISBN 0-9614088-5-5.

取自 “https://zh.wikipedia.org/w/index.php?title=奇异值分解&oldid=54716465”