Entangled states and the singular value decomposition

A. Gómez-Rodríguez
Instituto de Física, Universidad Nacional Autónoma de México
Apartado postal 20-364, México D.F., Mexico

Recibido el 17 de febrero 2000; aceptado el 21 de junio de 2000

In the first part of the paper we review the singular value decomposition (SVD). This tool is a generalization of the notion of diagonalization that can be used for matrices that can not be diagonalized in the usual sense. After stating the main theorem we discuss the use of the SVD to find the ranks of matrices and to approximate matrices (by other matrices of lesser rank). In the second part of the paper we show how the SVD can be used to assess if a given state in a tensor product space is entangled or not, and we provide an algorithm to produce the factors into which the state decomposes. The SVD can also provide a "best" (in the sense of the 2-norm) separable state corresponding to a given entangled state. The SVD provides a measure of the degree of entanglement. This measure is compared to other measures that have been proposed. As a by-product, we show that the SVD allows the numerical calculation of the so-called Schmidt decomposition of tensor product states. In the final part of the paper a possible extension of these methods to the case of statistical mixtures is proposed.

Keywords: Entangled states; singular value decomposition; Schmidt decomposition

En la primera parte de este trabajo revisamos la descomposición en valores singulares (SVD). Esta herramienta es una generalización de la noción de diagonalización que puede usarse con matrices que no pueden ser diagonalizadas en el sentido habitual. Después de mostrar el teorema principal discutimos el uso de la SVD para encontrar rangos de matrices y para aproximar matrices (por otras matrices de menor rango). En la segunda parte del trabajo mostramos cómo la SVD puede usarse para decidir si un estado dado en un espacio producto tensorial está enmarañado o no, y proporcionamos un algoritmo que da los factores en los que el estado se descompone. La SVD puede también dar "el mejor" (en el sentido de la 2-norma) estado separable que corresponde al estado enmarañado dado. La SVD da una medida del grado de enmarañamiento. Esta medida se compara con otras medidas que se han propuesto. Como producto colateral mostramos que la SVD permite el cálculo numérico de la llamada descomposición de Schmidt de estados producto tensorial. En la parte final del trabajo proponemos una extensión de estos métodos al caso de mezclas estadísticas.

Descriptores: Estados enredados; descomposición en estados singulares; descomposición de Schmidt

PACS: 02.10.sp; 03.65.Bz; 03.67.-a

1. Introduction

Consider two physical systems S_1 and S_2. According to the principles of quantum mechanics they are to be described by means of two Hilbert spaces H_1 and H_2. The compound system $S_3 = S_1 \cup S_2$ is also described by a third Hilbert space H_3 that is the tensor product of the spaces H_1 and H_2 so one writes

$$H_3 = H_1 \otimes H_2.$$

If a given vector $P \in H_3$ can be written as

$$p = a \otimes b$$

for some $a, b \in H_1 \otimes H_2$ then one says that the state p is unentangled (separable, factorable); otherwise the state is said to be entangled (non-separable).

There is a large (and growing) body of literature concerning the relevance of entanglement in several areas such as the transition from the quantum to the classical world, and quantum computing among others.

In this paper we present a simple criterion to decide whether a given state p is entangled or not. The criterion uses the so-called singular value decomposition, for this reason the first section is devoted to an elementary exposition of this tool given that it is not very well known in the physics community.

The singular value decomposition can also be used to provide a measure of the distance between a given state p and the nearest (in a suitable sense) unentangled state. In all cases our method provides the factors of the unentangled states involved.

This analysis of entanglement can be performed quite easily with the help of Maple, Matlab, Mathematica or any other numeric and/or symbolic software package.

The SVD also provides a simple numerical algorithm for performing the Schmidt decomposition of tensor product states.

2. The singular value decomposition (SVD)

In this section we present the singular value decomposition theorem in its various disguises and some of its applications such as the estimation of the rank of a matrix and matrix approximations. For a proof of the SVD theorem the reader is referred to Ref. 1.

Theorem 2.1 (Singular value decomposition, matrix form)

Given a matrix T (square or rectangular, real or complex) there exist matrices U, V and Σ such that:

1) $T = U\Sigma V^*$ (here V^* denotes the conjugate-transpose of V, or adjoint).

2) U and V are unitary (orthogonal in the real case).

3) Σ is real, diagonal, with non-negative elements arranged in decreasing order along the diagonal. It has the same size as T.

This factorization of T is called "singular value decomposition" or SVD for short.

The diagonal elements of Σ are called "singular values" of T. The columns of U are the "left singular vectors" of T whereas the columns of V are the "right singular vectors" of T.

Notice that

$$T^*T = V\Sigma^2V^*,$$

and

$$TT^* = U\Sigma^2U^*$$

so the right singular vectors of T are simply the eigenvectors of T^*T whereas the left singular vectors of T are the eigenvectors of TT^*. The singular values are the square roots of the eigenvalues of either T^*T or TT^*(with a few zeros added, as needed).

This theorem can be expressed in a somewhat longer form that is of relevance to our work:

Theorem 2.2 (Singular value decomposition)

Let V and U be inner product spaces over the field of complex or real numbers. Let $T : V \to U$ be a linear transformation, $\ker(T)$ the kernel of T, $\mathrm{Im}(T)$ the image of T, $r = \dim[\mathrm{Im}(T)]$ the rank of T. Let $m = \dim(V)$ and $n = \dim(U)$.

Then there exist an orthonormal basis

$$\{v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_m\}$$

of V, an orthonormal basis $\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ of U and real positive numbers $\sigma_1, \sigma_2, \ldots, \sigma_r$ such that:

1. $\{v_{r+1}, \ldots, v_m\}$ is an orthonormal basis of $\ker(T)$.
2. $\{u_1, u_2, \ldots, u_r\}$ is an orthonormal basis of $\mathrm{Im}(T)$.
3. $\{v_1, \ldots, v_r\}$ is an orthonormal basis of $[\ker(T)]^{\perp}$.
4. $\{u_{r+1}, \ldots, u_n\}$ is an orthonormal basis of $[\mathrm{Im}(T)]^{\perp}$.
5. $Tv_i = \sigma_i u_i$ and $T^*u_i = \sigma_i v_i$ for $i = 1, \ldots, r$.
6. $Tv_i = 0$ for $r < i \leq m$.
7. For any $w \in V$ $T(w) = \sum_{i=1}^{r} \sigma_i \langle v_i, w\rangle u_i$.

In Dirac's notation (favored by physicists, mostly when dealing with quantum mechanics) the SVD can be written simply as

$$T = \sum_{i=1}^{r} \sigma_i |u_i\rangle\langle v_i|,$$

and in this way it resembles a spectral decomposition or even a density matrix (this is nothing but the usual outer product expansion in a fancy notation).

This decomposition is useful mostly when (in the case $U = V$) the operator $T : V \to V$ is not normal (a normal operator is one that commutes with its adjoint so $[T, T^*] = 0$) because an operator is normal if and only if it has an orthonormal basis consisting of eigenvectors. When the operator T is normal it is usually more convenient to decompose (spectral decomposition) it as

$$T = \sum_{i=1}^{m} \lambda_i |v_i\rangle\langle v_i|,$$

where λ_i are the eigenvalues of T and $\{|v_i\rangle\}$ is an orthonormal basis of V consisting of eigenvectors of T ($T|v_i\rangle = \lambda_i|v_i\rangle$). However the real power of the SVD is manifest when $\dim(U) \neq \dim(V)$ and the matrix representing T with respect to a given basis is not square; unlike the spectral decompositions, the SVD remains valid for rectangular matrices.

Notice that the matrix of T with respect to the bases

$$\{v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_m\}$$

and

$$\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$$

is given by

$$M(T)_{i,j} = \langle u_i, Tv_j\rangle = \langle u_i, \sigma_j u_j\rangle = \sigma_j \delta_{i,j},$$

for $j = 1, 2 \ldots r$ and

$$M(T)_{i,j} = 0$$

otherwise.

Sometimes by singular values one means not just the set

$$\sigma_1, \sigma_2, \ldots, \sigma_r,$$

but the same set padded with zeros

$$\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0,$$

in order to complete p singular values with $p = \min\{m, n\}$.

2.1. Matrix norms

In the space of $n \times m$ matrices (or of linear transformations $T : V \to U$) we may define a norm (the 2-norm) by

$$\|T\| = \sup_{x \in V} \left(\frac{|Tx|}{|x|} \right) = \sup_{x \in V, |x|=1} (|Tx|),$$

(where $|Tx|$ refers to the vector norm in U and $|x|$ to the vector norm in V), and in terms of the SVD it can be shown that

$$\|T\| = \sigma_1$$

so the largest singular vector gives the matrix norm.

The so-called Frobenius norm (known as Hilbert-Schmidt norm in the physical literature)

$$\|T\|_F = \sqrt{\sum_{i,j} |T_{i,j}|^2} = \sqrt{\mathrm{Tr}\,(TT^*)}$$

can also be expressed in terms of the singular values as

$$\|T\|_F = \sqrt{\sum_i |\sigma_i|^2}.$$

2.1.1. *Matrix approximations*

As we have seen

$$T = \sum_{i=1}^{r} \sigma_i |u_i\rangle\langle v_i|,$$

so we may wonder about the "goodness" of approximations given by

$$T_h = \sum_{i=1}^{h} \sigma_i |u_i\rangle\langle v_i|,$$

with $h < r$. Notice that T_h is a matrix of rank h.

It can be shown that [2]

$$\|T - T_h\| = \sigma_{h+1}$$

and the SVD gives a quantitative measure of how good the approximation is.

Another measure of the goodness of the approximation is the relative distance defined as

$$Rd = \frac{\|T - T_h\|}{\|T\|} = \frac{\sigma_{h+1}}{\sigma_1}$$

that satisfies $0 \le Rd \le 1$.

Actually T_h is the "best" rank h approximation to T, in the sense that it achieves the minimum of $\{\|T - T_0\| \,|T_0$ has rank $h\}$.

2.1.2. *Rank of a matrix*

In linear algebra courses the lecturers cheat¡. When they teach us how to compute the rank of a matrix they always use "nice" matrices such as

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

whose rank can be determined in many ways (for instance by means of reduction to the Hermite Normal Form). But what about the matrix?

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10^{-7} & 0 \\ 0 & 0 & 10^{-45} \end{pmatrix}$$

which, strictly speaking has rank 3. If your experimental error is of the order of, say, 10^{-20} then certainly you may (you must¡) consider the 10^{-45} term as zero and conclude that rank $(T) = 2$. If you are using a calculator with six significant figures then the rank has to be taken as one¡.

The SVD is the only tool (to our knowledge) that can provide a sensible analysis of the rank for "real life" matrices. In principle the rank of a matrix is the number of singular values different from zero. In practice you must choose a tolerance tol such that any singular value less than tol will be considered as zero. tol is normally chosen on the basis of the experimental error in the measurements, in any case tol is larger than the smallest number that the computer can meaningfully consider as different from zero [Matlab uses $MAX(SIZE(T)) * NORM(T) * EPS$, where EPS is 2^{-52}].

In machines with IEEE arithmetic, tol should be larger than the smallest number larger than one that can be represented in the computer. For a PC it will be of the order of 10^{-16}.

3. On entanglement

Consider the spaces $\mathbb{C}^n$ and $\mathbb{C}^m$ and define their tensor product by

$$a \otimes b = ab^T = |a\rangle \otimes |b\rangle$$

(a and b are $n \times 1$ and $m \times 1$ matrices, respectively) so it is an element of $\mathbb{C}^{n \times m}$ (the space of all complex-valued n by m matrices). This is the tensor product we will use, all other tensor products are isomorphic to this one. The corresponding tensor product space will be written accordingly as

$$\mathbb{C}^{n \times m} = \mathbb{C}^n \otimes \mathbb{C}^m.$$

A vector of $|v\rangle$ in $\mathbb{C}^{n \times m}$ is called "factorable" (or "separable") if it can be written as

$$|v\rangle = |a\rangle \otimes |b\rangle$$

with $|a\rangle \in \mathbb{C}^n$ and $|b\rangle \in \mathbb{C}^m$. Otherwise it is called "entangled".

Next think of $p \in \mathbb{C}^{n\times m}$ as a linear transformation

$$p : \mathbb{C}^m \to \mathbb{C}^n$$
$$v \mapsto pv$$

Theorem: $p = |a\rangle \otimes |b\rangle$ for some $|a\rangle \neq 0 \in \mathbb{C}^n$ and $|b\rangle \neq 0 \in \mathbb{C}^m$ if and only if rank $(p) = 1$.

Proof: If $p = |a\rangle \otimes |b\rangle$ then $pv = a(b^T v)$. In this case Im (p) = span (a) and rank $(p) = 1$

Conversely, if rank $(p) = 1$ then by the SVD theorem p can be written as $\sigma ab^* = \sigma a(\bar{b})^T$ where the vectors a and b are normalized.

This theorem gives a practical way to determine if a given state is entangled or not: construct the tensor product matrix and determine its rank; if the rank is one then the state is factorable, otherwise it is entangled. In case the state is factorable, the SVD yields the factors (the left and the right singular vector corresponding to the non-zero singular value).

But there remains a problem: as explained above, to determine whether the rank of p is one or not can be a tricky business if the matrix is pathological. For this reason we propose a slightly different strategy. Given a matrix $p \in \mathbb{C}^{n\times m}$ and its SVD, say $[U, \Sigma, V] = svd(p)$, then we can form the best (in the 2-norm sense explained above) rank-one approximation to p as

$$p1 = \Sigma(1,1) * U(:,1) * V(:,1)'$$

(here we used Matlab's colon notation: $U(:,1)$ is the first column of U and $V(:,1)$ is the first column of V, $V(:,1)'$ is the conjugate transpose of $V(:,1)$) and $\Sigma(2,2)$ is the distance (in the 2-norm again) between p and $p1$.

Then if $a = \Sigma(1,1) * U(:,1)$ and $b = \overline{V(:,1)}$ $p1 = a \otimes b = ab^T$

A useful Matlab (or Maple) program would be, then, one giving $\Sigma(1,1)$, $\Sigma(2,2)$, $U(:,1)$ and $V(:,1)$.

For instance, take p as

$$p = \begin{pmatrix} 1 & 2 & 3 \\ 1.001 & 2.001 & 3.001 \end{pmatrix},$$

then

$$U = \begin{pmatrix} -.70696 & -.70726 \\ -.70726 & .70696 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} 5.2926 & 0 & 0 \\ 0 & 4.6281 \times 10^{-4} & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} -.34 & .87285 & .40825 \\ -.53454 & .21817 & -.8165 \\ -.80175 & -.43651 & .40825 \end{pmatrix}.$$

Notice that the procedure does not require the matrix to be square.

This example illustrates the main features of the proposed algorithm. First notice that the input matrix p was designed as a matrix with two rows almost identical (second row is first row plus the vector $[.01, .01, .01]$). This fact is reflected in the fact that the second singular value is quite small but not zero.

As a second example consider the maximally entangled state

$$|a\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

(here both H_1 and H_2 are two-dimensional) so the relevant matrix is

$$m = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix},$$

with singular values $\sigma_1 = \sigma_2 = 1/\sqrt{2}$. Thus $|a\rangle$ is at a distance $1/\sqrt{2}$ of the closest unentangled state, the relative distance achieves the maximum value of $Rd = \sigma_2/\sigma_1 = 1$; the closest state is $|a\rangle_c = \frac{1}{\sqrt{2}}|00\rangle$ (actually, in this case, $\frac{1}{\sqrt{2}}|11\rangle$ is at the same distance).

For all these reasons it is concluded that σ_2 (or σ_2/σ_1) gives a measure of the degree of entanglement: it tells if the state is entangled or not and it also gives the distance to the closest unentangled state. In addition, it provides a link to modern linear algebra (via the SVD).

3.1. General spaces

Let U and V be finite-dimensional vector spaces over the field of complex numbers and let $n = \dim(U)$ $m = \dim(V)$. Let $W = U \otimes V$ be the tensor product space of U and V.

If we introduce bases for these spaces

$$\alpha = \{u_1, u_2, \ldots u_n\},$$
$$\beta = \{v_1, v_2, \ldots v_m\},$$
$$\gamma = \{u_i \otimes v_j\},$$

then, given $a \in U$ and $b \in V$ we can write

$$a = \sum_{i=1}^{n} a_i u_i,$$
$$b = \sum_{j=1}^{m} b_j v_j,$$
$$a \otimes b = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j u_i \otimes v_j,$$

and proceed as in the previous section using the coordinates of a, b and $a \otimes b$ which are in $\mathbb{C}^n$, $\mathbb{C}^m$ and $\mathbb{C}^{n\times m} = \mathbb{C}^n \otimes \mathbb{C}^m$ respectively. In other words, there was no loss of generality in the treatment of the previous section since U is isomorphic to $\mathbb{C}^n$, V is isomorphic to $\mathbb{C}^m$ and $U \otimes V$ is isomorphic to $\mathbb{C}^n \otimes \mathbb{C}^m$.

3.2. The Schmidt decomposition

As a by-product of the technique described above we obtain the so-called Schmidt decomposition of a state in $H_1 \otimes H_2$. Let $p \in H_1 \otimes H_2$, then decomposing p by means of the SVD we obtain

$$p = \sum_{i=1}^{r} \sigma_i |u_i\rangle\langle v_i| = \sum_{i=1}^{r} \sigma_i U^i V^{i*}$$

(here U^i and V^i are the columns of U and V respectively). Defining $W^i = \text{conj}\,(V^i)$ we have that

$$p = \sum_{i=1}^{r} \sigma_i U^i W^{iT}$$

and

$$p = \sum_{i=1}^{r} \sigma_i U^i \otimes W^i,$$

where the σ_i are real and non-negative (for this reason in the literature one frequently finds $\sigma' = \sigma^2$ and $p = \sum_{i=1}^{r} \sqrt{\sigma_i'} U^i \otimes W^i$)

3.3. Statistical mixtures

When the states in $H_3 = H_1 \otimes H_2$ are not pure states and are described by a density operator ρ the SVD algorithm for disentanglement can still be applied with some modifications. Below we sketch the procedure, the details are presented in a forthcoming publication:

- By means of a spectral decomposition expand the density matrix ρ as

$$\rho = \sum_{i=1}^{n} \lambda_i |u_i\rangle\langle u_i|$$

  where n is the dimension of the tensor product space, $\{\lambda_i\}$ are the eigenvalues of ρ and $\{|u_i\rangle\}$ is an orthonormal basis of the space formed by eigenvectors of ρ.

- Next use the SVD algorithm to analyze each $|u_i\rangle \in H_1 \otimes H_2$. The measure of entanglement in this case would be the set of the second singular values for each $|u_i\rangle$.

3.3.1. *When are all these methods likely to be useful?*

In the literature, most of the time, the problems are solved exactly in algebraic closed form. When numbers are needed (for instance to exemplify the techniques) simple whole numbers or fractions are used. But in real life data are obtained from experiments!.

For instance, if the density matrix ρ is needed, it can be obtained by measuring various observables $\{A_i\}$ and since

$$\langle A_i \rangle = \text{Tr}\,(\rho A_i)$$

(by a suitable choice of the set $\{A_i\}$) ρ can be completely determined [5]. But then the numbers to be used in the measurement of entanglement are subject to experimental errors and numerical techniques such as the SVD are absolutely necessary. The SVD provides the tools to deal with experimental (and also numerical) error so decisions (concerning rank, for instance) can be made.

4. Conclusions

In this work we have achieved the following:

1. We have shown that a state is unentangled if and only if the rank of the tensor product matrix is one.

2. We have shown how to determine the rank of the tensor product matrix. We argue that (given numerical errors or even experimental ones) the right tool is the SVD.

3. We have shown how to find the distance (in the 2-norm) from a given state in $H_1 \otimes H_2$ to the nearest (in the 2-norm) unentangled state.

4. We have provided an algorithm such that given $p \in H_1 \otimes H_2$ it gives:

   - The first singular value of p. This number is also the 2-norm of p.
   - The second singular value of p. This number is zero if and only if the rank of p is 1 and if and only if the state is unentangled. This number is also the 2-norm distance between p and the unentangled state $p1 \in H_1 \otimes H_2$ that is closest to p.
   - It provides the unentangled state $p1 \in H_1 \otimes H_2$ that is closest to p. referred to above.
   - It gives the factors of p; *i.e.* it gives the vectors a and b such that $p1 = a \otimes b$

   Using the second singular value σ_2 as a measure of entanglement has several distinct advantages such as:

   - $\sigma_2 = 0$ if the state is disentangled
   - Is invariant under unitary similarity transformations (*i.e.* does not depend on the choice of basis)

The two properties above are considered as requirements for any "good" entanglement measure (see Ref. 3 and references therein). Our measure probably satisfies the third requirement as formulated by Ref. 3, since a related measure (Frobenius norm) does indeed satisfy it [4] and in $\mathbb{C}^n$ all norms are equivalent. Certainly σ_2 does not reflect the "information content" unlike the familiar measures of the form $\sum_{i=1}^{n} \text{Tr}\,(\rho_i \ln(\rho_i))$.

Certainly the SVD is not the only tool available for the determination of the rank of a matrix. Other possible choices include the Hermite Normal Form (or row-reduced echelon form) and the QR (orthogonal-triangular Householder) decomposition. However the SVD is more robust and by far the best choice if the matrices are suspected to be close to rank-deficiency.

It has been shown that the SVD can be used to perform the Schmidt decomposition of a tensor product state.

All the results have been stated in terms of the 2-norm, however they remain valid, with minor modifications, in the Frobenius norm $\|T\|_F = \sqrt{\sum_i |\sigma_i|^2} = \sqrt{\mathrm{Tr}\,(TT^*)}$ that gives rise to the so-called Hilbert-Schmidt metric. The measure of entanglement proposed by Witte and Trucks [4] is based precisely on this norm; in our language their measure is simply $\sqrt{\sum_{i=2} |\sigma_i|^2}$ where the σ_i are, as before, the singular values.

Finally a possible extension of these methods to the case of statistical mixtures has been proposed.

No attempt has been made to discuss the infinite-dimensional case.

Acknowledgments

The author acknowledges support from CONACYT (grant 0088P-E) and DGAPA-UNAM (grant IN107296).

1. G.H. Golub and C.F. Van Loan, *Matrix Computations*, (The Johns Hopkins University Press, Baltimore and London, 1989).
2. B.J. Noble and W. Daniel, *Applied Linear Algebra*, (Prentice-Hall, Englewood Cliffs, NJ, 1977).
3. V. Vedral and M.B. Plenio, *Phys. Rev.* **57** (1998) 1619.
4. C. Witte and M. Trucks, Preprint submitted to Elsevier, (april 1999).
5. K. Blum, *Density Matrix Theory and Applications*. (Plenum, New York, 1981).