(μ) Schmidt Decomposition

In this mathematical digression, we introduce the Schmidt decomposition of a vector in a tensor product of Hermitian spaces $\mathcal{H}_A \otimes \mathcal{H}_B$, which is equivalent to the singular value decomposition of a matrix of a linear map $\mathcal{H}_B \to \mathcal{H}_A$.

Let us fix a vector $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, written in the standard basis $|ab\rangle = |a\rangle \otimes |b\rangle$ in the form

$$|\Psi\rangle = \sum_{a \in A, b \in B} \alpha_{a,b} |a\rangle \otimes |b\rangle.$$

Assume, for instance, that $d_A = \dim A \leqslant d_B = \dim B$ (otherwise the argument is similar after exchanging the roles of A and B), and write it in the form

$$|\Psi\rangle = \sum_{a \in A} |a\rangle \otimes |\phi_a\rangle$$

where $|\phi_a\rangle = (\langle a| \otimes \mathbb{1}_B)|\Psi\rangle = \sum_b \alpha_{a,b}|b\rangle \in \mathcal{H}_B$.

We want to orthogonalize the vectors $|\phi_a\rangle$ by choosing a suitable basis of $\mathcal{H}_A$. Then, once the nonzero vectors among them have been normalized, the vectors $\frac{|\phi_a\rangle}{\|\phi_a\|}$ will be orthonormal vectors of $\mathcal{H}_B$, and in the two new orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$, only the d_A diagonal coefficients $\|\phi_a\| \geqslant 0$ (indexed by $a \in A$) will appear in the decomposition of $|\Psi\rangle$.

To evaluate the inner products $\langle \phi_a | \phi_{a'} \rangle = \operatorname{Tr} |\phi_{a'} \rangle\langle \phi_a|$, let us compute the partial trace over B of the positive operator $|\Psi\rangle\langle\Psi|$:

$$\operatorname{Tr}_B(|\Psi\rangle\langle\Psi|) = \operatorname{Tr}_B\Big(\sum_{a,a' \in A} |a'\rangle\langle a| \otimes |\phi_{a'}\rangle\langle\phi_a|\Big) = \sum_{a,a' \in A} \langle\phi_a|\phi_{a'}\rangle\, |a'\rangle\langle a|.$$

This operator is positive because $|\Psi\rangle\langle\Psi| \geqslant 0 \implies \operatorname{Tr}_B(|\Psi\rangle\langle\Psi|) \geqslant 0$, and it determines all these inner products $\langle\phi_a|\phi_{a'}\rangle$. If we choose a basis $|a\rangle = |a\rangle_A$ of $\mathcal{H}_A$ consisting of eigenvectors of the operator $\operatorname{Tr}_B(|\Psi\rangle\langle\Psi|)$, we have the spectral decomposition

$$\operatorname{Tr}_B(|\Psi\rangle\langle\Psi|) = \sum_{a \in A} \lambda_a |a\rangle\langle a|.$$

Then, by identifying coefficients, we find $\langle\phi_a|\phi_{a'}\rangle = 0$ for $a \neq a'$; the $|\phi_a\rangle$ are indeed orthogonal, with $\langle\phi_a|\phi_a\rangle = \|\phi_a\|^2 = \lambda_a \geqslant 0$. After normalizing the nonzero ones by writing $|a\rangle_B = \frac{|\phi_a\rangle}{\|\phi_a\|} = \frac{|\phi_a\rangle}{\sqrt{\lambda_a}} \in \mathcal{H}_B$ when $\lambda_a > 0$ (and choosing arbitrary orthonormal completion vectors when $\lambda_a = 0$), we obtain the decomposition

$$|\Psi\rangle = \sum_{a \in A} |a\rangle_A \otimes |\phi_a\rangle = \sum_{a \in A} \sqrt{\lambda_a}\, |a\rangle_A |a\rangle_B,$$

where $|\phi_a\rangle = \sqrt{\lambda_a}\,|a\rangle_B$. Setting $\sigma_a = \sqrt{\lambda_a}$, and if necessary extending the $|a\rangle_B$ to an orthonormal basis of $\mathcal{H}_B$ when $d_B > d_A$ by the incomplete basis theorem, we have proved the following theorem.

Theorem (Schmidt decomposition). *Given a vector $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $d_A \leqslant d_B$, there exist two orthonormal bases $|a\rangle_A$ of $\mathcal{H}_A$ and $|b\rangle_B$ of $\mathcal{H}_B$ in which $|\Psi\rangle$ decomposes in diagonal form*

$$|\Psi\rangle = \sum_{a\in A} \sigma_a\, |a\rangle_A |a\rangle_B,$$

where the coefficients $\sigma_a \geqslant 0$ are called the ***Schmidt coefficients*** *or* ***singular values*** *of $|\Psi\rangle$. These Schmidt coefficients $\sigma_a = \sqrt{\lambda_a}$ are uniquely determined by $|\Psi\rangle$ as the square roots of the eigenvalues λ_a of $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)$.*

If $d_A \geqslant d_B$, the same argument (exchanging the roles of A and B) gives the decomposition

$$|\Psi\rangle = \sum_{b\in B} \sigma_b\, |b\rangle_A |b\rangle_B,$$

where the Schmidt coefficients $\sigma_b = \sqrt{\lambda_b}$ are uniquely determined by $|\Psi\rangle$ as the square roots of the eigenvalues λ_b of $\mathrm{Tr}_A(|\Psi\rangle\langle\Psi|)$. We can combine the two cases by writing, for example, $|i\rangle_A$ ($i = 1, 2, \ldots, d_A$) for the orthonormal basis of $\mathcal{H}_A$, and $|j\rangle_B$ ($j = 1, 2, \ldots, d_B$) for the orthonormal basis of $\mathcal{H}_B$. The Schmidt decomposition is written

$$|\Psi\rangle = \sum_{i} \sigma_i\, |i\rangle_A |i\rangle_B,$$

where the sum ranges over $i = 1, \ldots, \min(d_A, d_B)$ and the Schmidt coefficients are uniquely determined by $|\Psi\rangle$. With this notation, note that one has both $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = \sum_i \sigma_i^2 |i\rangle\langle i|_A$ and $\mathrm{Tr}_A(|\Psi\rangle\langle\Psi|) = \sum_i \sigma_i^2 |i\rangle\langle i|_B$, so these two operators have the **same nonzero eigenvalues** $\lambda_i = \sigma_i^2$.

The Schmidt decomposition can equivalently be viewed for an arbitrary **matrix** M of a linear map $\mathcal{H}_B \to \mathcal{H}_A$, which is written in the form

$$M = \sum_{a\in A, b\in B} M_{a,b} |a\rangle\langle b|.$$

If, with the outer product $|a\rangle\langle b|$, one associates the corresponding tensor product $|a\rangle \otimes |b\rangle$, this defines a linear isomorphism called **vectorization**, where M is transformed (by stacking its rows into a single column vector) into the vector

$$|\Psi\rangle = \sum_{a\in A, b\in B} M_{a,b} |ab\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

Thus the matrix M and the vector $|\Psi\rangle$ may be regarded as two equivalent representations that determine the same coefficients $M_{a,b}$.

The Schmidt decomposition of the vector $|\Psi\rangle$ therefore means that there exist two orthonormal bases $|a\rangle_A$ of $\mathcal{H}_A$ and $|b\rangle_B$ of $\mathcal{H}_B$ in which the matrix (a rectangular matrix of size $d_A \times d_B$) decomposes in diagonal form

$$M = \sum_{a\in A} \sigma_a\, |a\rangle_A \langle a|_B \quad \text{or} \quad M = \sum_{b\in B} \sigma_b\, |b\rangle_A \langle b|_B,$$

depending on whether $d_A \leqslant d_B$ or $d_A \geqslant d_B$. We therefore have a matrix decomposition:

Theorem (Singular value decomposition). *Given a matrix $M = (M_{a,b})$ of a linear map $\mathcal{H}_B \to \mathcal{H}_A$, there exist two orthonormal bases $|i\rangle$ of $\mathcal{H}_A$ and $|j\rangle$ of $\mathcal{H}_B$ in which M is written in diagonal form*

$$M = \sum_i \sigma_i \, |i\rangle_A \langle i|_B,$$

where the sum ranges over $i = 1, \ldots, \min(d_A, d_B)$ and the coefficients $\sigma_i \geqslant 0$ are the ***singular values*** *of M.*

If U denotes the unitary matrix whose columns are the vectors $|u_i\rangle = |i\rangle$ in the standard basis of the $|a\rangle$, and V the unitary matrix whose columns are the vectors $|v_j\rangle = |j\rangle$ in the standard basis of the $|b\rangle$, the coefficients of M in the new bases are $\langle u_a|M|v_b\rangle = \langle Ua|M|Vb\rangle = \langle a|U^*MV|b\rangle$. Consequently, the singular value decomposition theorem for M amounts to saying that there exist two unitary matrices U (of size $d_A \times d_A$) and V (of size $d_B \times d_B$) such that

$$U^*MV = \Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & 0 \\ & & \ddots & \\ & 0 & & \ddots \\ & & & & \sigma_\delta \end{pmatrix}$$

where $\delta = \min(d_A, d_B)$; the matrix Σ is in general rectangular, of size $d_A \times d_B$, with the singular values σ_i as its diagonal entries. Since $U^{-1} = U^*$ and $V^{-1} = V^*$, we can invert the relation:

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

which is equivalent to the singular value decomposition $M = \sum_i \sigma_i |u_i\rangle\langle v_i|$.

Remark. When $d_A \leqslant d_B$, the operator $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = \sum_{a,a' \in A} \langle\phi_a|\phi_{a'}\rangle \, |a'\rangle\langle a|$ can be computed in terms of $M = (\alpha_{a,b})$. Since $|\phi_a\rangle = \sum_b \alpha_{a,b}|b\rangle$, we have

$$\langle\phi_a|\phi_{a'}\rangle = \sum_b \alpha_{a',b}\alpha^*_{a,b} = (MM^*)_{a',a}.$$

Thus, identifying $\sum_{a,a'} \langle\phi_a|\phi_{a'}\rangle \, |a'\rangle\langle a| = \sum_{a,a'} (MM^*)_{a',a}|a'\rangle\langle a|$, we obtain

$$\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = MM^*.$$

We can check directly that the singular value decomposition $M = U\Sigma V^*$ gives $MM^* = U\Sigma V^*V\Sigma^*U^* = U\Sigma\Sigma^*U^* = U\Lambda U^*$, where Λ is the square $d_A \times d_A$ diagonal matrix with diagonal entries $\lambda_i = \sigma_i^2$, which are indeed the eigenvalues of $\mathrm{Tr}_B(|\Psi\rangle\langle\Psi|) = MM^*$ in the spectral decomposition of $MM^* = U\Lambda U^*$.

Similarly, if $d_A \geqslant d_B$, we find $\mathrm{Tr}_A(|\Psi\rangle\langle\Psi|) = M^*M$, whose spectral decomposition is $M^*M = V\Sigma^*\Sigma V^* = V\Lambda' V^*$, where Λ' is the square $d_B \times d_B$ diagonal matrix with diagonal entries $\lambda_i = \sigma_i^2$.

Since the rank of the matrix M does not depend on the choice of bases in the expression of M, it is equal to the number of nonzero singular values σ_i. Returning to $|\Psi\rangle$, this gives the following definition.

Definition (Schmidt rank). *The rank r of a vector $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is the number of its nonzero singular values (counted with multiplicity). It is equal to the rank of the matrix M.*

We can therefore write $|\Psi\rangle = \sum_i \sigma_i |i\rangle_A |i\rangle_B = \sum_i \sigma_i |u_i\rangle|v_i\rangle$, where the sum contains r nonzero terms. If $r = 1$, $|\Psi\rangle$ can be written as a tensor product $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. Conversely, if $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, we can always complete $|\psi_A\rangle$ (normalized) to an orthonormal basis of $\mathcal{H}_A$ and $|\psi_B\rangle$ (normalized) to an orthonormal basis of $\mathcal{H}_B$, hence obtaining a Schmidt decomposition of rank 1.

Theorem (Characterization of an entangled state). *A state $|\Psi\rangle$ is a product state if and only if its Schmidt rank is equal to 1. Consequently, a state $|\Psi\rangle$ is* ***entangled*** *if and only if its Schmidt rank is > 1.*

In the special case $d_A = d_B$, M is a square matrix representing an operator, and Σ is a positive diagonal square matrix. The singular value decomposition $M = U\Sigma V^*$ implies $MM^* = U\Sigma V^* V \Sigma U^* = U\Sigma^2 U^*$ and, similarly, $M^*M = V\Sigma^2 V^*$. If M is not normal, these two positive matrices MM^* and M^*M are not equal: they have the same eigenvalues σ_i^2, but with different bases of eigenvectors (the columns of U and of V, respectively). By analogy with the definition of the absolute value $|M| = \sqrt{M^*M} = \sqrt{MM^*}$ in the case where M is normal, one can nevertheless always define in the general case the **two absolute values**:

$$|M|' = \sqrt{MM^*} = U\Sigma U^* \quad \text{and} \quad |M| = \sqrt{M^*M} = V\Sigma V^*.$$

The singular value decomposition $M = U\Sigma V^*$ can then be interpreted as either of the two decompositions

$$M = U\Sigma U^*(UV^*) = |M|'W,$$
$$M = (UV^*)V\Sigma V^* = W|M|,$$

where $W = UV^*$ is unitary, since $WW^* = UV^*VU^* = UU^* = \mathbb{1}$. Writing $W = e^{i\Theta}$, where Θ is Hermitian, gives the following decomposition.

Theorem (Polar decomposition). *Every square matrix M representing an operator decomposes in the form*

$$M = |M|' e^{i\Theta} \qquad \text{or} \qquad M = e^{i\Theta}|M|,$$

where the matrices $|M|' = \sqrt{MM^*}$ and $|M| = \sqrt{M^*M}$ are positive and Θ is Hermitian.

In dimension 1, this is simply the polar decomposition $m = |m|e^{i\theta}$ of $m \in \mathbb{C}$.

Example. Given the POVM $(E_m)_m$, a corresponding generalized measurement is defined by matrices K_m satisfying $K_m^* K_m = E_m$, with an unnormalized post-measurement state $K_m|\psi\rangle$. The general form of the solutions K_m of the equation $K_m^* K_m = E_m$ is given by their polar decomposition $K_m = U\sqrt{E_m}$, where U is unitary. Indeed,

$$K_m^* K_m = \sqrt{E_m}\, U^* U \sqrt{E_m} = \sqrt{E_m}\, \mathbb{1} \sqrt{E_m} = (\sqrt{E_m})^2 = E_m.$$

The unnormalized post-measurement state is then $U\sqrt{E_m}|\psi\rangle$.