## CSE439 Week 13: Matrix Algebra and the SVD

What happens if we try to take the "inner product" of two $m \times n$ matrices $A$ and $B$ by first "unrolling" them as vectors? Remembering to conjugate the entries of $A$, we get

$$\langle A, B \rangle \;=\; \sum_{\substack{i=1,m \\ j=1,n}} \overline{A[i,j]} B[i,j].$$

Now let $C = A^* B$. Since $A^*$ is $n \times m$, this is an $n \times n$ square matrix. From

$$C[r,s] = \sum_{k=1}^{m} A^*[r,k] B[k,s] \;=\; \sum_{k=1}^{m} \overline{A[k,r]} B[k,s]$$

we get that the diagonal entries of $C$ are $C[r,r] = \sum_{k=1}^{m} \overline{A[k,r]} B[k,r]$. Hence the diagonal sum gives

$$\sum_{r=1}^{n} C[r,r] \;=\; \sum_{\substack{k=1,m \\ r=1,n}} \overline{A[k,r]} B[k,r] \;=\; \langle A, B \rangle$$

as we defined it above. The diagonal sum at left is called the **trace**, with notation $Tr(C)$. Now for a vector $v$, the self inner-product $\langle v, v \rangle$ gives the squared Euclidean norm of $v$, written $||v||_2^2$, so $||v||_2 \;=\; \sqrt{\langle v, v \rangle}$. The analogous concept for matrices is the **Frobenius norm**, named for Ferdinand Georg Frobenius:

$$||A||_F \;=\; \sqrt{Tr\left(A^* A\right)}.$$

Or you can simply say it's the Euclidean 2-norm of the vector obtained by "unrolling" the matrix. This norm, however, overstates the *action* of the matrix in Euclidean space, which involves its $m \times n$ dimensions. This is

$$||A||_2 \;=\; \sup\{||Av||_2 : v \text{ is a unit vector of length } n\}.$$

For some further remarks: Since our vectors are finite-dimensional, the "ball surface" of unit vectors is **compact**, which actually *means* that there is a definite vector $v$ that maximizes $||Av||_2$ rather than just having a limit---so we can write "max" in place of "sup" for "supremum." The task of *finding* such a vector $v$ is the main algorithmic need of computing the **singular value decomposition** (**SVD**) as treated below. It tumbles out of the SVD Theorem that $||A||_2 \;\leq\; ||A||_F$ for every matrix $A$. But the inutition is that $||A||_2$ tells the most that $A$ can "stretch" a vector along the fixed dimensions it operates on, whereas $||A||_F$ is the maximum amount of "stretch" that the entries of $A$ could give under any configuration of dimensions.

## Trace, Density Matrices, and Measurements

When we do $|\phi\rangle\langle\phi|$ for a quantum state vector $\phi = [a_1, \ldots, a_N]^T$, the diagonal entries $a_i\overline{a_i}$ of the outerproduct give $||\phi||_2^2$, which equals $1$ since $\phi$ is a unit vector. Since a **density matrix** $\rho$ is a linear combination of outerproducts $|\phi\rangle\langle\phi|$ by weights summing to $1$, the trace $Tr(\rho)$ is also 1. Now when a unitary matrix $U$ acts on $\phi$, the density matrix of the resulting vector $U\phi$ is

$$|U\phi\rangle\langle U\phi| \;=\; |U\phi\rangle\langle\phi|U^* \;=\; U|\phi\rangle\langle\phi|U^*.$$

By linear additivity, a unitary operator acts on a mixed state $\rho$ by the double action $U\rho U^*$. Put all this together, and the rule is that *the trace of a density matrix is always* $1$. The action by unitary matrices preserves the trace. Ultimately this is just the idea of probabilities summing to $1$.

These ideas play into the most general idea of measurement on which there is wide consensus. It generalizes the notion of a projective measurement of a pure state. Recall that a **positive semidefinite** (**PSD**) matrix is one of the form $M^*M$ for some matrix $M$.

**Definition**: A **positive operator valued measure** (**POVM**) is a set $\{E_1, \ldots, E_m\}$ of PSD matrices such that $\sum_{j=1}^{m} E_j \;=\; \mathbf{I}$ (text has $E_j^*E_j$ there, is it a typo?). Given a mixed state $\rho$ the probability $p_j$ of outcome $j$ is given by

$$p_j \;=\; Tr(E_j\rho).$$

If a PSD representation $E_j \;=\; M_j^*M_j$ is specified for each $j$ (it might not be unique, but specifying it is part of the measurement apparatus) then the next state is

$$\rho' \;=\; \frac{M_j\rho M_j^*}{p_j}.$$

We can use this to answer a natural question: How does the mixed state $\frac{1}{2}\big(|0\rangle + |1\rangle\big)$ differ from the quantum superposition $|+\rangle \;=\; \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)$? Besides the different constant, there is a difference in meaning that dictates that when mixed states are involved, we really need to use the density matrix representation of both. So we are really talking about $\frac{1}{2}\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big)$ versus $|+\rangle\langle +|$. We have that the former is

$$\frac{1}{2}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right) \;=\; \begin{bmatrix}0.5 & 0\\ 0 & 0.5\end{bmatrix}$$

while the latter is

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \;=\; \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

Both matrices have trace $1$; the difference is that $\mathbf{J} = |+\rangle\langle +|$ hasd non-zero off-diagonal elements. Also $\mathbf{J}^2 = \mathbf{J}$, which is the definition of when a density matrix represents a pure state. Now we know that $\mathbf{H}|+\rangle = |0\rangle$, which the density matrix under the double-action rogers:

$$\mathbf{H}\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}\mathbf{H} \;=\; \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \;=\; \frac{1}{4}\begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \;=\; \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \;=\; |0\rangle\langle 0|,$$

whereas

$$\mathbf{H}\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}\mathbf{H} \;=\; \frac{1}{2}\mathbf{H}\mathbf{H} \;=\; \frac{1}{2}\mathbf{I} \;=\; \frac{1}{2}\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big)$$

back again. So whereas Alice would measure $0$ with certainty if she applied Hadamard to $|+\rangle$, when she does so to her mixed state she will still get $0$ with only 50% probability. The kicker is that if she instead measures in the $\{|+\rangle, |-\rangle\}$ basis, whether before or after applying the Hadamard gate, she will get those outcomes with 50% probability each. Thus "a mixed state does not remember which pure states were used to define it." The only reality it has---at least the only reality that we can *know*---is its density matrix.

## Traceout and Spectral Purification

[I have decided to insert this material after all.]

A further rule involving density matrices and tensor products starts from pure states $|\phi\rangle$ and $|\psi\rangle$. Recall that the adjoint $(|\phi\rangle \otimes |\psi\rangle)^*$ is $\big(|\phi\rangle^* \otimes |\psi\rangle^*\big) \;=\; \langle\phi| \otimes \langle\psi|$. That is, we don't reverse the product as we would with ordinary matrix multiplication. The indexing is "tiered" in the form $state(xy)$ where $x$ pertains to the space of $|\phi\rangle$ ("Alice") and $y$ to the space of $|\psi\rangle$ ("Bob"). So now involving outerproducts and running indices $u$ over Alice's row space and $v$ over Bob's:

$$\big(|\phi\rangle \otimes |\psi\rangle\big)\big(\langle\phi| \otimes \langle\psi|\big)[uv, xy] \;=\; \big(|\phi\rangle \otimes |\psi\rangle\big)[uv]\big(\langle\phi| \otimes \langle\psi|\big)[xy] \;= \phi(u)\psi(v)\overline{\phi(x)}\,\overline{\psi(y)}.$$

Whereas,

$$\Big(\big(|\phi\rangle\langle\phi|\big) \otimes \big(|\psi\rangle\langle\psi|\big)\Big)[uv, xy] \;=\; \big(|\phi\rangle\langle\phi|\big)[ux]\big(|\psi\rangle\langle\psi|\big)[vy] \;=\; \phi(u)\overline{\phi(x)}\psi(v)\overline{\psi(y)},$$

which is the same. Note that the left-hand side of the second equation is the tensor product of two pure

density matrices. By additive linearity for tensor products, this proves the general rule:

> The density matrix of two unentangled systems can be represented as the tensor product of density matrices of the respective systems. In symbols: $\rho_{A,B} = \rho_A \otimes \rho_B$. (Here we understand identity up to multiplication by unit scalars.)

A nifty point is that we can semi-invert this process *even when Alice and Bob are entangled*. The operation is called the **traceout**. It is easiest to picture and execute when we apply it to the second tier of the whole space, i.e., in *"tracing out Bob."* It is also called the **partial trace** $Tr_{\mathbb{B}}$ mapping elements of the "higher space" $\mathbb{A} \otimes \mathbb{B}$ to the space $\mathbb{A}$. Given the density matrix $\rho$ of the whole system:

- Block out $\rho$ into square submatrices as-if it were a tensor product $A \otimes B$. If Bob holds $k$ qubits, then the submatrices will be $2^k \times 2^k$.
- Replace each submatrix by its trace. When you consider the submatrices on the main diagonal, you can see the overall trace is unchanged---it is still $1$ as it must be for $\rho$.
- The resulting matrix is the density $\rho_A$ for Alice "after tracing out Bob."

There is also a matrx $\rho_B$ of Bob "tracing out Alice." However, it need not follow that $\rho = \rho_A \otimes \rho_B$. That happens if (and only if? the things that occur to you on the second pass...) Alice and Bob were initially unentangled. In that case, *all* of the "Bob" submatrices have trace $1$. The effect is the same---in the case above where Alice and Bob are pure states---as substituting $\psi(v) = \psi(y) = 1$. This leaves $\phi(u)\overline{\phi(x)}$, which is the $ux$ entry of Alice's outerproduct $|\phi\rangle\langle\phi|$.

**Example 1**: The traceout of the entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is done by

$$Tr_B\left(\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix},$$

which is the density matrix of the completely mixed state again. Notice incidentally that $\frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle)$ gives exactly the same pure-state vector $\frac{1}{\sqrt{2}}[1,0,0,1]^T$, even before we take its outerproduct to get the above $4 \times 4$ density matrix. This all reinforces that Alice applying $\mathbf{H}$ or whatever unitary operation on her half of the entangled pair has no effect on the current state of her knowledge of it, which is represented by the density matrix.

**Example 2**: The pure state $\frac{1}{2}(|000\rangle + |001\rangle + |110\rangle - |111\rangle)$ from the "practice Prelim II" has the following density matrix, with the big-endian indexing included:

$$\frac{1}{4}\left[\begin{array}{c|cccccccc} & 000 & 001 & 010 & 011 & 100 & 101 & 110 & -111 \\ \hline 000 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 001 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 010 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 011 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 101 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 110 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ -111 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \end{array}\right]$$

Notice that the two off-diagonal traces cancel. So tracing out the third qubit leaves:

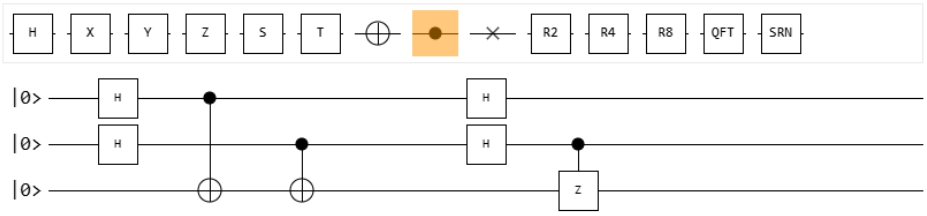
$$\frac{1}{4}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}.$$

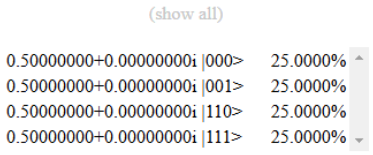
This is not the density matrix of an entangled pair. Nor is it even the completely mixed state on two qubits. It is the mixture $0.5\big(|00\rangle\langle 00| + |11\rangle\langle 11|\big)$. Its tensor product with Bob holding $|+\rangle$ is not the same as the above density matrix---the corners stay zeroed out. Thus the fact that the above pure state is entangled carries through here as well.

Whereas, if we use a **CCZ** gate---or just a **CZ** gate on the second and third qubits---then the minus sign flips to make $\frac{1}{2}\big(|000\rangle + |001\rangle + |110\rangle + |111\rangle\big)$. Now the pure density matrix is

$$\frac{1}{4}\left[\begin{array}{c|cccccccc} & 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ \hline 000 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 001 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 010 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 011 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 101 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 110 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 111 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Tracing out Bob does leave the density matrix $\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ of the pure state $\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$ in Alice's hands. And the whole system truly is $(\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big) \otimes \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)$. Here is the quantum circuit, including the final **CZ** gate:




The interpretation that might go with this is that the first two Hadamard and CNOT gates tried to entangle qubits 1 and 3 and then entangled 2 and 3. The hope was that by "entangling 3 twice" we could actually *disentangle* it and leave 1 and 2 entangled. This didn't quite work, but it does work if we apply the $\mathbf{CZ}$ gate after the second Hadamard gate. The subject of **entanglement swapping** commonly needs 6 qubits to illustrate, but this gives some of the flavor at smaller scale.

One more notable fact:

**Theorem**: For every mixed state $\rho$ on $n$ qubits there is a pure state $|\kappa\rangle$ on $2n$ qubits such that tracing out the last $n$ qubits in $|\kappa\rangle\langle\kappa|$ leaves exactly $\rho$.

**Proof**: Because $\rho$ is Hermitian, we can find an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_N$ (where $N = 2^n$ as usual) and real eigenvalues $\lambda_1, \ldots, \lambda_N$ such that $\rho = \lambda_1|\mathbf{u}_1\rangle\langle\mathbf{u}_1| + \cdots + \lambda_N|\mathbf{u}_N\rangle\langle\mathbf{u}_N|$. This gives a diagonal matrix in the coordinates of the eigenbasis, but we can apply a unitary change of basis $U$ to make $U\rho U^*$ diagonal in the standard basis. Since we've seen that this double-action preserves the trace, which is $1$ in $\rho$, we get $\lambda_1 + \cdots + \lambda_N = 1$. (Note how this says $\rho$ is far different from a unitary matrix, even a unitary matrix that is Hermitian, because those have each individual eigenvalue being of magnitude $1$.) Now define

$$\kappa = \sqrt{\lambda_1}(|\mathbf{u}_1\rangle \otimes |\mathbf{u}_1\rangle) + \cdots + \sqrt{\lambda_N}\big(|\mathbf{u}_N\rangle \otimes |\mathbf{u}_N\rangle\big).$$

This is a legal pure state because the squares of the amplitudes $\sqrt{\lambda_i}$ sum to $1$. So let us apply the traceout to $|\kappa\rangle\langle\kappa|$.

When we do $|\kappa\rangle\langle\kappa|$, we get cross-terms but they stay within each $n$-qubit tier of the whole Hilbert space---by the point we observed at the beginning of this section. Within each tier, they have the form $|\mathbf{u}_i\rangle\langle\mathbf{u}_j|$ with $i \neq j$ (multiplied by $\sqrt{\lambda_i\lambda_j}$). Now the main diagonal of this outerproduct is $\sum_{k=1}^{N} \mathbf{u}_i(k)\overline{\mathbf{u}_j(k)}$, which is exactly the *inner* product $\langle\mathbf{u}_j|\mathbf{u}_i\rangle$. This in turn is zero because $\mathbf{u}_i$ and $\mathbf{u}_j$ are orthogonal. So taking the trace of these ``Bob'' submatrices makes the off-diagonal components of the traceout vanish without a trace. The only survivors are the terms

$$\left|\sqrt{\lambda_i}\mathbf{u}_i \otimes \mathbf{u}_i\right\rangle\left\langle\sqrt{\lambda_i}\mathbf{u}_i \otimes \mathbf{u}_i\right| = \lambda_i|\mathbf{u}_i \otimes \mathbf{u}_i\rangle\langle\mathbf{u}_i \otimes \mathbf{u}_i|.$$

Now tracing out "Bob" in these submatrices just substitutes $1$ for the second $\mathbf{u}_i$, leaving

$$\lambda_1|\mathbf{u}_1\rangle\langle\mathbf{u}_1| + \cdots + \lambda_N|\mathbf{u}_N\rangle\langle\mathbf{u}_N|,$$

which is the original $\rho$ back again. ☒

There are other possible pure states on higher numbers of qubits that can do the same. In Example 2 above, we saw that the 2-qubit mixture $0.5\big(|00\rangle\langle 00| + |11\rangle\langle 11|\big)$ is the traceout of the 3-qubit pure state $\frac{1}{2}\big(|000\rangle + |001\rangle + |110\rangle - |111\rangle\big)$, where we got the off-diagonal cancellations without needing to go to a full-blown spectral representation. Moreover, the Bell pair $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ is *exactly* the $|\kappa\rangle$ from the completely mixed state $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$, which is already in spectral form. The general name for this process is **mixed-state purification**. It often happens that the neatest way to calculate or prove results about mixed states is to "lift" them to pure states in a larger space, calculate in the higher space, and then trace back down. John Smolin of IBM T.J. Watson gave this technique the evocative name of "appealing to the church of the higher Hilbert Space."

## Quantum Ontology and Epistemology

*Ontology* has to do with *being*; *epistemology* with *knowing*. We have taken the view that pure states "are"---that is, they have existence unto themselves. We represent them as state vectors, but at exponentially high cost in many cases. For properly mixed states $\rho$, this is less clear. We can regard some pure state $|\kappa\rangle$ from a higher space that traces out to it as its ontology, but (a) $|\kappa\rangle$ is far from unique, and (b) as indicated by the use of "$N$" in the last proof, it often comes at exponential cost.

The epistemological side, however, has given a remarkably consistent set of answers for over a century now:

- The only way we can gain knowledge about a quantum state, whether pure or mixed, is by measuring it.
- All measurements of a pure state $|\phi\rangle$ go---explicitly or implicitly---through its density matrix $\rho_\phi = |\phi\rangle\langle\phi|$.
- Operations on density matrices "gibe" with measurements and probabilities in ways already prescribed by (Bayesian!) conditional inferencing.
- All scalar quantities involved in this reckoning are *real* numbers denoting (conditional) *probabilities*, not "amplitudes".

The last point is part of why Richard Lipton and I have mused about giving an account of quantum reality without complex numbers. For the above, where Hermitian not unitary matrices are primary, complex numbers need only be seen in components of orthonormal eigenvectors $\mathbf{u}_i$, such as

$\mathbf{u}_1 = \frac{1}{\sqrt{2}}[1, i]^T$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}[i, 1]^T$ on both the practice and actual Prelim II. For reality, however, the notion of *phase* seems inescapable, and complex numbers (IMPHO) give its best treatment. There are wide indications that Bloch spheres---in higher dimensions as well as for single qubits---are physically real. They give a description via two real numbers $\theta$ and $\varphi$; note that $\cos(\theta)$ is a probability, not an amplitude. However, $\varphi$ is a phase angle and governs whether and where complex numbers appear in other figuring. So *Reali*ty strikes back but doesn't completely subjugate the complex realm, which is necessitated by the Fundamental Theorem of Algebra anyway.

This finally leaves the super-skeptical question of whether there is a bedrock of *being* beneath what is knowable. One form of this question is whether the notion of an *observer*---often styled as a *conscious* observer---is essential to existence. This idea long predates quantum mechanics. It was formulated as philosophical *subjective idealism* by the Irish Anglican bishop George Berkeley in the early 1700s, whom the city of Berkeley in California is expressly named after. It is well captured by the following pair of limericks---my mod of what Fr. Ronald Knox wrote two centuries later:

A divinity student said, "God
Must find it exceedingly odd
That the Warden's plum tree
Continues to be
When there's no one about in the Quad."

And the reply as Knox imagined in a newspaper's Letters column:

"Dear Sir, your perplexity's odd.
*I* am always about in the Quad.
So the Warden's plum tree
Shall continue to be,
Since observed by---
    yours faithfully, ---God."

Whether the advent of quantum mechanics enhances such arguments over the pervasive presence of an unseen benevolent God is not something I choose to amplify. Lipton and I broadly sympathize with Samuel Johnson's reply of refuting Berkeley by kicking a stone---fully aware that the kick involves the exertion of quantum mechanical electric force on a surface whose solidity is effected by vibrating molecules. Speaking for myself as a Christian, I hold a halfway position toward *fideism* that disclaims logical proof and reproducible knowledge of God, and I regard this as merely orthodox. There is, however, considerable reason to assert the pervasive presence of an unseen---and vaguely malign(?)---"Bob" in the form of entanglements with unknown systems, even going back to the Big Bang. Entanglements with outside nature, developed both now and prior, are the current best explanation for *decoherence*. The above illustrations of Tom Brady-style "deflation" in off-diagonal parts of density matrices are well representative of decoherence.

Staying completely with "Nature's Rose", we will conclude with the matter of classical simulation of quantum algorithms, via advanced computational methods. First on the list is the Singular Value Decomposition, which is the closest an arbitrary---not even square---matrix can come to the blessings of both unitary and Hermitian properties.

## The SVD

A matrix $S$ is (**pseudo-**)diagonal if it is (**non-**)square and $S[i,j] = 0$ whenever $i \neq j$. It follows that both $S^*S$ and $SS^*$ are diagonal square matrices. Some of the diagonal entries may be $0$.

**SVD Theorem**: For every $m \times n$ matrix $A$ we can efficiently find:

- an $m \times m$ unitary matrix $U$,
- an $m \times n$ pseudo-diagonal matrix $\Sigma$ with non-negative entries $\Sigma[i,i] = \sigma_i$, and
- an $n \times n$ unitary matrix $V$,

such that $A = U\Sigma V^*$. Furthermore, we can arrange that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)}$, and in consequence:

- $||A||_F = \sqrt{\sum_i \sigma_i^2}$.
- $||A||_2 = \sigma_1$,
- $A^*A = V\Sigma^T U^* U\Sigma V^* = V \operatorname{diag}\left(\sigma_i^2\right) V^*$, and
- $AA^* = U\Sigma V^* V\Sigma^T U^* = U \operatorname{diag}\left(\sigma_i^2\right) U^*$,

so that the squares of the $\sigma_i$ and associated vectors give the spectral decompositions of the Hermitian PSD matrices $A^*A$ and $AA^*$, respectively.

The $\sigma_i$ are the **singular values**. The number $r$ of positive ones equals the **rank** of $A$. Whereas some of the $\lambda_i$ can be negative in the Spectral Theorem---when the Hermitian matrix is not PSD---none of the $\sigma_i$ is negative. The first $r$ columns of $U$ form an orthonormal basis for the subspace $\mathbb{W}$ spanned by the columns of $A$ (called the column space of $A$), while the first $r$ columns of $V$ form an orthonormal basis for the column space of $A^*$. The latter is identical to the row space of $A$ when $A$ is a real matrix---and in that case, $U$ and $V$ come out being real as well. The remaining $m - r$ columns of $U$ form an orthonormal basis for the space $\mathbb{W}^\perp$, which is also the **nullspace** of $A^*$. As with the Spectral Theorem, the basis vectors are not unique when there is multiplicity or when we don't have $r = m = n$, but the values $\sigma_i$ are unique (when sorted in nonascending order, so we can say the matrix $\Sigma$ is unique too). Once $U$ and $V$ are specified, we get $\Sigma = U^*AV$ too.

**Proof**: The procedure works by recursion through subspaces and so resembles the proof of the Spectral Theorem. The first and top-level step is most emblematic. It begins by *finding* a unit vector $v_1$ that maximizes $||Av_1||_2$. Then $\sigma_1 = ||Av_1||_2$ is the first and biggest singular value. It can't be zero (unless $A$ is the all-zero matrix, in which case we've "hit triviality"), so

$$u_1 \;=\; \frac{Av_1}{\sigma_1}$$

is a unit vector. If there are more than one maximizing unit vectors $v$ then we will get multiplicity, but let us first suppose that the $v_1$ and associated $u_1$ are unique. Before doing the recursion, we may postulate that $u_1$ is arbitrarily extended to an orthonormal basis $U_1$ of $\mathbb{C}^m$ (or of $\mathbb{R}^m$ in the real case) and $v_1$ to an orthonormal basis $V_1$ of $\mathbb{C}^n$. In the resulting coordinates, we get

$$U_1^* A V_1 \;=\; \begin{bmatrix} \sigma_1 & w_1^* \\ 0 & B \end{bmatrix} \;=\; S_1$$

for some vector $w_1$ of length $n-1$ and $(m-1)\times(n-1)$ matrix $B$. The red $0$ stands for $m-1$ zeroes and is because $Av_1 = \sigma_1 u_1$ so there is no dependence on the other $m-1$ coordinates. *The goal is to prove that $w_1$ must be all-zero too.* Then recursing on $B$ will hammer out the (pseudo-)diagonal matrix $\Sigma$.

Let $w = \begin{bmatrix} \sigma_1 \\ w_1 \end{bmatrix}$ as a column vector. Then $S_1 w \;=\; w' = \begin{bmatrix} \sigma_1^2 + w_1^* w_1 \\ B w_1 \end{bmatrix}$. Ignoring the $Bw$ part, we get $||w'|| \;\geq\; \sigma_1^2 + w_1^* w_1$. The right-hand side equals $||w||^2$. Dividing by $||w||_2$ hence gives us

$$\frac{||S_1 w||_2}{||w||_2} \;\geq\; ||w||_2 \;=\; \sqrt{\sigma_1^2 \;+\; w_1^* w}.$$

Now if $w_1$ is nonzero, then $w_1^* w_1$ is a positive real number, so $\dfrac{||S_1 w||_2}{||w||_2} \;>\; \sigma_1$. Under the definition of the 2-norm for matrices, this means $||S_1||_2 \;>\; \sigma_1$. But

$$||S_1||_2 \;=\; ||U_1^* A V_1||_2 \;=\; ||A||_2$$

because $U_1$ and $V_1$ are unitary. And $||A_2|| \;=\; \sigma_1$ by how we defined $\sigma_1$. This is a contradiction saying "$\sigma_1 > \sigma_1$." The only way out is for $w_1$ to be a zero vector.

The recursion then takes place on the perpendicular subspace of $v_1$, or in general, the perpendicular subspace of the span of the orthogonal unit vectors $v_j$ chosen thus far. The final point is that the corresponding vectors $u_j$ also come out orthogonal. This is because, when $i \neq j$ (and at stages where $\sigma_i$ and $\sigma_j$ are both nonzero---else we are in the base case of completing orthonormal bases on the nullspaces):

$$\sigma_i \sigma_j \langle u_i | u_j \rangle \;=\; \langle \sigma_i u_i | \sigma_j u_j \rangle \;=\; \langle Av_i | Av_j \rangle \;=\; v_i^* A^* A v_j \;=\; v_i^* \sigma_j^2 v_j \;=\; \sigma_j^2 \langle v_i | v_j \rangle \;=\; 0,$$

finally using the orthogonality of the $v_i$ vectors. The fourth equality happens because $v_j$ is an eigenvector of $A^*A$ with eigenvalue $\sigma_j^2$. The reason given by the (short!) proof in the MIT notes (https://math.mit.edu/classes/18.095/2016IAP/lec2/SVD_Notes.pdf) is that

$$A^*A \;=\; \left(U\Sigma V^*\right)^*\left(U\Sigma V^*\right) \;=\; V\Sigma^T U^* U\Sigma V^* \;=\; V\Sigma^T\Sigma V^*,$$

which in turn converts to the way we have been writing the spectral decomposition since $V$ is unitary. However, substituting $U^*U = I$ strikes me as assuming what one is trying to prove about the $u_i$ vectors.

To tie up the loose end, *we choose to restart the proof*. We apply the original Spectral Theorem to the Hermitian PSD matrix $A^*A$ to get nonnegative eigenvalues $\lambda_1, \ldots, \lambda_n$---listed in nonincreasing order---and orthonormal eigenvectors $v_1, \ldots, v_n$ such that

$$A^*A \;=\; \lambda_1|v_1\rangle\langle v_1| \;+\; \cdots \;+\; \lambda_n|v_n\rangle\langle v_n| \;=\; V^*\mathrm{diag}(\lambda_i)V,$$

taking $V$ as the matrix with the eigenvectors as its columns. Now define $\sigma_i$ to be the nonnegative square root of $\lambda_i$ for each $i$. Since the rank $r$ of $A$ equals the rank of $A^*A$, we get $\sigma_i > 0$ for $i = 1$ to $r$. For these $i$, define

$$u_i \;=\; \frac{Av_i}{\sigma_i}.$$

*Now* the above demonstration that $\langle u_i | u_j \rangle = 0$ is logically valid, because we arranged that $\sigma_i^2 = \lambda_i$ is an eigenvalue of $A^*A$ with eigenvector $v_i$ in advance. What we've lost, however, is the original proof's definition of $\sigma_i$ so that $u_i$ is a unit vector. We recover it, however, this way:

$$\langle u_i | u_i \rangle = \left\langle \frac{Av_i}{\sigma_i} \middle| \frac{Av_i}{\sigma_i} \right\rangle = \frac{1}{\sigma_i^2} v_i^* A^* A v_i \;=\; \frac{1}{\sigma_i^2} v_i^* \lambda_i v_i \;=\; v_i^* v_i \;=\; 1.$$

And $u_i$ is an eigenvector of $AA^*$ because

$$AA^* u_i = AA^* \frac{Av_i}{\sigma_i} \;=\; \frac{1}{\sigma_i} A\left(A^*A\right)v_i \;=\; \frac{1}{\sigma_i} A\sigma_i^2 v_i \;=\; \sigma_i A v_i \;=\; \sigma_i^2 u_i.$$

For $i > r$, we can arbitrarily complete the basis by choosing orthonormal vectors that span the nullspace.

So now the only thing we've "lost" compared to the first proof strategy is the fact that at the first and each later step of the recursion, the choice of unit vector $v_i$ maximizes $||Av_i||_2$. However, now we can

appeal to the uniqueness of the $\lambda_i$ and "quasi-uniqueness" of the eigenvectors up to the flex of multiplicity. The squares of the $\sigma_i$ and the $\lambda_i$ **must** coincide. What comes out is a deep fact that the largest eigenvalues of $A^*A$ naturally pick out the directions in which $A$ stretches the most. ☒

**Corollary**: For a square matrix $A$ already of the form $E^*E$ (and that goes for any Hermitian PSD matrix), the SVD and spectrum of $E$ coincide with $U = V$.

**Proof**. The diagonal form $E = U\Lambda U^*$ has the specified properties; because $E$ is PSD, the $\lambda_i$ are nonnegative, and we can arrange $U$ so that the diagonal is in nonincreasing order. ☒

In all other cases where $A$ is diagonalizable, there are reasons for saying the SVD gives *more* information than the diagonalization. This is especially so with upper or lower traingular matrices---see example below. And of course, there are many square matrices that can't be diagonalized...to say nothing of non-square matrices...for which the SVD is the only game in town.

Our two-pronged proof suggests two different algorithms for *computing* the SVD of a matrix $A$:

- Diagonalize $A^*A$ to get $\lambda_i$'s and $V$, then $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{Av_i}{\sigma_i}$.
- Find a unit vector $v$ maximizing $||Av||_2$ and recurse.

Other methods come into play when $A$ has certain particular features. Niloufer Mackey developed new methods in her 1993 UB CS PhD dissertation under Patricia Eberlein.

## Examples and Applications

In any upper or lower triangular matrix $A$, the elements of the diagonal *are* the eigenvalues. They are thus independent of all the off-diagonal entries at upper right. Those entries have information that does get picked up by the SVD. The two examples in the MIT notes are good for this.

**Example 1**:

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$

$$A^* = A^T = \begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix}$$

$$A^*A = \begin{bmatrix} 9+16 & 20 \\ 20 & 25 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \quad AA^* = \begin{bmatrix} 9 & 12 \\ 12 & 16+25 \end{bmatrix} = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$$

**Example 2**:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ \epsilon? & 0 & 0 & 0 \end{bmatrix}$$