Lecture 9 – ECE 275A

Singular Value Decomposition (SVD)
Eigenstructure of $A^H A$

Let $A : \mathcal{X} = \mathbb{C}^n \rightarrow \mathcal{Y} = \mathbb{C}^m$ be an $m \times n$ matrix operator mapping between two Cartesian complex Hilbert spaces.

Recall that (with $A^H = A^*$ for $A$ a mapping between Cartesian spaces)

$$r(AA^H) = r(A) = r(A^H) = r(A^H A)$$

Therefore the number of nonzero (and hence strictly positive) eigenvalues of $AA^H : \mathbb{C}^m \rightarrow \mathbb{C}^m$ and $A^H A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ must both be equal to $r = r(A)$.

Let the nonnegative eigenvalues of $A^H A$ be denoted and ordered as

$$\sigma_1^2 \geq \cdots \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2 = 0$$

with corresponding $n$-dimensional orthonormal eigenvectors

$$\begin{align*}
\text{span of } \mathcal{R}(A^H) &= \mathcal{N}(A)^\perp \\
\text{spans } \mathcal{N}(A) &= \mathcal{N}(A^H A)
\end{align*}$$
Eigenstructure of $A^H A$ – Cont.

Thus we have

$$(A^H A) v_i = \sigma_i^2 v_i \quad \text{with} \quad \sigma_i^2 > 0 \quad \text{for} \quad i = 1, \cdots, r$$

and

$$(A^H A) v_i = 0 \quad \text{for} \quad i = r + 1, \cdots, n$$

The eigenvectors $v_{r+1} \cdots v_n$ can be chosen to be any orthonormal set spanning $\mathcal{N}(A)$.

An eigenvectors $v_i$ associated with a distinct nonzero eigenvalues $\sigma_i^2$, $1 \leq i \leq r$, is unique up to sign $v_i \mapsto \pm v_i$.

Eigenvectors $v_i$ associated with the same nondistinct nonzero eigenvalue $\sigma_i^2$ with multiplicity $p$ can be chosen to be any orthonormal set that spans the $p$-dimensional eigenspace associated with that eigenvalue.

Thus we see that there is a lack of uniqueness in the eigen-decomposition of $A^H A$. This lack of uniqueness (as we shall see) will carry over to a related lack of uniqueness in the SVD.

What is unique are the values of the nonzero eigenvalues, the eigenspaces associated with those eigenvalues, and any projection operators we construct from the eigenvectors (uniqueness of projection operators).
Eigenstructure of $A^H A$ – Cont.

In particular, we uniquely have

$$P_{\mathcal{R}(A^H)} = V_1 V_1^H \quad \text{and} \quad P_{\mathcal{N}(A)} = V_2 V_2^H$$

where

$$V_1 \triangleq \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix} \in \mathbb{C}^{n \times r} \quad V_2 \triangleq \begin{pmatrix} v_{r+1} & \cdots & v_n \end{pmatrix} \in \mathbb{C}^{n \times \nu}$$

and

$$V \triangleq \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{pmatrix} \in \mathbb{C}^{n \times n}$$

Note that

$$\mathcal{R}(V_1) = \mathcal{R}(A^H), \quad \mathcal{R}(V_2) = \mathcal{N}(A), \quad \mathcal{R}(V) = \mathcal{X} = \mathbb{C}^n$$

$$I_{n \times n} = V^H V = V V^H = V_1 V_1^H + V_2 V_2^H = P_{\mathcal{R}(A^H)} + P_{\mathcal{N}(A)}$$

$$I_{r \times r} = V_1^H V_1, \quad I_{\nu \times \nu} = V_2^H V_2$$

It is straightforward to show that $V_1 V_1^H$ and $V_2 V_2^H$ are idempotent and self-adjoint.
Eigenstructure of $A^H A$ – Cont.

We now prove two identities that will prove useful when deriving the SVD.

Taking $\sigma_i = \sqrt{\sigma_i^2}$ define

$$S_{r \times r} \triangleq \text{diag}(\sigma_1 \cdots \sigma_r)$$

Then

$$A^H A v_i = \sigma_i^2 v_i \quad 1 \leq i \leq r$$

can be written as

$$A^H A V_1 = V_1 S^2$$

which yields

$$I_{r \times r} = S^{-1} V_1^H A^H A V_1 S^{-1} \tag{1}$$

We also note that

$$A^H A v_i = 0 \iff A v_i \in \mathcal{R}(A) \cap \mathcal{N}(A^H) = \{0\}$$

so that $A^H A v_i = 0, i = r + 1, \cdots, n$ yields

$$0_{m \times \nu} = A V_2 \tag{2}$$
Eigenstructure of $AA^H$

The eigenstructure of $A^H A$ determined above places constraints on the eigenstructure of $AA^H$.

Above we have shown that

$$(A^H A)v_i = \sigma_i^2 v_i \quad i = 1, \ldots, r$$

where $\sigma_i^2$, $1 \leq i \leq r$, are nonzero. If we multiply both sides of this equation by $A$ we get (recall that $r = r(AA^H) \leq m$

$$(AA^H)(Av_i) = \sigma_i^2 (Av_i) \quad i = 1, \ldots, r$$

Showing that $Av_i$ and $\sigma_i^2$ are eigenvector-eigenvalue pairs.

Since $AA^H$ is hermitian, the vectors $Av_i$ must be orthogonal. In fact, the vectors

$$u_i \triangleq \frac{1}{\sigma_i} Av_i \quad 1 \leq i \leq r$$

are orthonormal.
This follows from defining

\[ U_1 = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix} \in \mathbb{C}^{m \times r} \]

which is equivalent to

\[ U_1 = AV_1S^{-1} \]

and noting that Equation (1) yields orthogonality of the columns of \( U_1 \)

\[ U_1^H U_1 = S^{-1}V_1^H A^H A V_1 S^{-1} = I \]

Note from the above that

\[ S_{r \times r} = U_1^H A V_1 \quad (3) \]

Also note that a determination of \( V_1 \) (based on a resolution of the ambiguities described above) completely specifies \( U_1 = AV_1 S^{-1} \). Contrawise, it can be shown that a specification of \( U_1 \) provides a unique determination of \( V_1 \).
Eigenstructure of $AA^H$ - Cont.

Because $u_i$ correspond to the nonzero eigenvalues of $AA^H$ they must span $\mathcal{R}(AA^H) = \mathcal{R}(A)$. Therefore

$$\mathcal{R}(U_1) = \mathcal{R}(A) \quad \text{and} \quad P_{\mathcal{R}(A)} = U_1 U_1^H$$

Complete the set $u_i$, $i = 1, \cdots, r$, to include a set of orthonormal vectors, $u_i$, $i = r + 1, \cdots m$, orthogonal to $\mathcal{R}(U_1)$ (this can be done via random selection of new vectors in $\mathbb{C}^m$ followed by Gram-Schmidt orthonormalization.) Let

$$U_2 = \begin{pmatrix} u_{r+1} & \cdots & u_m \end{pmatrix}_{m \times \mu}$$

with $\mu = m - r$.

By construction

$$\mathcal{R}(U_2) = \mathcal{R}(U_1) = \mathcal{R}(A) = \mathcal{N}(A^H)$$

and therefore

$$0 = A^H U_2$$

and

$$P_{\mathcal{N}(A^H)} = U_2 U_2^H$$
Eigenstructure of $AA^H$ – Cont.

Setting

$$U = \begin{pmatrix} u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{pmatrix} = \begin{pmatrix} U_1 & U_2 \end{pmatrix}$$

we have

$$I_{m \times m} = U^H U = UU^H = U_1 U_1^H + U_2 U_2^H = P_R(A) + P_N(A^H)$$
Derivation of the SVD

\[ \Sigma_{m \times n} \triangleq U^H AV = \begin{pmatrix} U_1^H \\ U_2^H \end{pmatrix} A \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} U_1^H AV_1 & U_1 AV_2 \\ U_2^H AV_1 & U_2^H AV_2 \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \]

or

\[ A = U \Sigma V^H \]

Note that

\[ A = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^H \\ V_2^H \end{pmatrix} = U_1 SV_1^H \]

This yields the Singular Value Decomposition (SVD) factorization of \( A \)

\[ \text{SVD: } A = U \Sigma V^H = U_1 SV_1^H \]

Note that when \( A \) is square and full rank, we have \( U = U_1, V = V_1, \Sigma = S \), and

\[ A = USV^H \]
**SVD Properties**

The matrix $U$ is unitary, $U^{-1} = U^H$, and its columns form an orthonormal basis for the codomain $\mathcal{Y} = \mathbb{C}^m$ wrt to the standard inner product. (Its rows are also orthonormal.) If we denote the columns of $U$, known as the left singular vectors, by $u_i$, $i = 1, \cdots, m$, the first $r$ lsv’s comprise the columns of the $m \times r$ matrix $U_1$, while the remaining lsv’s comprise the columns of the $m \times \mu$ matrix $U_2$, where $\mu = m - r$ is the dimension of the nullspace of $A^*$. The lsv $u_i$ is in one-to-one correspondence with the singular value $\sigma_i$ for $i = 1, \cdots, r$.

The matrix $V$ is unitary, $V^{-1} = U^H$, and its columns form an orthonormal basis for the domain $\mathcal{X} = \mathbb{C}^n$ wrt to the standard inner product. (Its rows are also orthonormal.) If we denote the columns of $V$, known as the right singular vectors, by $v_i$, $i = 1, \cdots, m$, the first $r$ rsv’s comprise the columns of the $n \times r$ matrix $V_1$, while the remaining rsv’s comprise the columns of the $n \times \nu$ matrix $V_2$, where $\nu = n - r$ is the nullity (dimension of the nullspace) of $A$. The rsv $v_i$ is in one-to-one correspondence with the singular value $\sigma_i$ for $i = 1, \cdots, r$.

We have

$$A = U_1 S V_1^H = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} v_1^H \\ \vdots \\ v_r^H \end{pmatrix} = \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H$$

Each term in this dyadic expansion is unique (i.e., does not depend on how the ambiguities mentioned above are resolved).
We can use the SVD to gain geometric intuition of the action of the matrix operator $A : \mathcal{X} = \mathbb{C}^n \rightarrow \mathcal{Y} = \mathbb{C}^m$ on the space $\mathcal{X} = \mathcal{R}(A^H) + \mathcal{N}(A)$.

The action of $A$ on $\mathcal{N}(A) = \mathcal{R}(V_2)$ is trivial to understand from its action on the right singular vectors which form a basis for $\mathcal{N}(A)$,

$$A v_i = 0 \quad i = r + 1, \ldots, n$$

In class we discussed the geometric interpretation of the action of the operator $A$ on $\mathcal{R}(A^H)$ based on the dyadic expansion

$$A = \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H$$

as a mapping of a hypersphere in $\mathcal{R}(A^H)$ to an associated hyperellipsoid in $\mathcal{R}(A)$ induced by the basis vector mappings

$$v_i \xrightarrow{A} \sigma_i u_i \quad i = 1, \ldots, r$$
**SVD Properties – Cont.**

When $A$ is square and presumably full rank, $r = n$, this allows us to measure the numerical conditioning of $A$ via the quantity (the *condition number* of $A$)

$$\text{cond}(A) = \frac{\sigma_1}{\sigma_n}$$

This measures the degree of ‘flattening’ (distortion) of the hypersphere induced by the mapping $A$. A perfectly conditioned matrix $A$ has $\text{cond}(A) = 1$, and an infinitely ill-conditioned matrix has $\text{cond}(A) = +\infty$.

Using the fact that for square matrices $\det A = \det A^T$ and $\det AB = \det A \det B$, we note that

$$1 = \det I = \det UU^H = \overline{\det U} \det U = |\det U|^2$$

or

$$|\det U| = 1$$

and similarly for the unitary matrices $U^H$, $V$, and $V^H$. (Note BTW that this implies for a unitary matrix $U$, $\det U = e^{i\phi}$ for some $\phi \in \mathbb{R}$. When $U$ is real and orthogonal, $U^{-1} = U^T$, this reduces to $\det U = \pm 1$.) Thus for a square matrix $A$,

$$|\det A| = \det USU^H = |\det U| \cdot |\det S| \cdot |\det U| = \det S = \sigma_1 \sigma_1 \cdots \sigma_n$$
Exploiting identities provable from the M-P Theorem (see Lecture 7 and Homework 3) we have

\[ A^+ = V \Sigma^+ U^H \]

\[ = (V_1 \; V_2) \begin{pmatrix} S^+ & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_1^H & U_2^H \end{pmatrix} \]

\[ = V_1 S^{-1} U_1^H \]

\[ = \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} \\ \vdots \\ \frac{1}{\sigma_r} \end{pmatrix} \begin{pmatrix} u_1^H \\ \vdots \\ u_r^H \end{pmatrix} \]

\[ = \frac{1}{\sigma_1} v_1 u_1^H + \cdots + \frac{1}{\sigma_r} v_r u_r^H \]

Note that this construction works regardless of the value of the rank \( r \). This shows that knowledge of the SVD of a matrix \( A \) allows us to determine its p-inv even in the rank-deficient case. Also note that the pseudoinverse is \textit{unique}, regardless of the particular SVD variant (i.e., it does not depend on how the ambiguities mentioned above are resolved).
**SVD Properties – Cont.**

Note that having an SVD factorization of $A$ at hand provides us with an orthonormal basis for $\mathcal{X} = \mathbb{C}^n$ (the columns of $V$), an orthonormal basis for $\mathcal{R}(A^H)$ (the columns of $V_1$), an orthonormal basis for $\mathcal{N}(A)$ (the columns of $V_2$), an orthonormal basis for $\mathcal{Y} = \mathbb{C}^m$ (the columns of $U$), an orthonormal basis for $\mathcal{R}(A)$ (the columns of $U_1$), and an orthonormal basis for $\mathcal{N}(A^H)$ (the columns of $U_2$).

Although the SVD factorization, the bases mentioned above, are not uniquely defined, it is the case that the orthogonal projectors constructed from the basis are unique (from uniqueness of projection operators). Thus *we can construct the unique orthogonal projection operators via*

\[
P_{\mathcal{R}(A)} = U_1 U_1^H \quad P_{\mathcal{N}(A^H)} = U_2 U_2^H \quad P_{\mathcal{R}(A^H)} = V_1 V_1^H \quad P_{\mathcal{N}(A)} = V_2 V_2^H
\]

Obviously having access to the SVD is tremendously useful. With the background we have now covered, one can now can greatly appreciate the utility of the Matlab command `svd(A)` which returns the singular values, left singular vectors, and right singular vectors of $A$, from which one can construct all of the entities described above. (Note that the singular vectors returned by Matlab will not necessarily *all* agree with the ones you construct by other means because of the ambiguities mentioned above. However, *the singular values will be the same*, and the left and right singular vector associated with the same, *distinct* singular value should only differ from yours by a sign at most.) Another useful Matlab command is `pinv(A)` which returns the pseudoinverse of $A$ regardless of the value of the rank of $A$.  

---

ECE 275AB Lecture 9 – Fall 2008 – V1.1 – © K. Kreutz-Delgado, UC San Diego – p. 15/17
Two Simple SVD Examples

In the third homework assignment you are asked to produce the SVD for some simple matrices by hand and then construct the four projection operators for each matrix as well as the pseudoinverse.

The problems in Homework 3 have been carefully designed so that you do not have to perform eigendecompositions to obtain the SVD’s. Rather, you can easily force the matrices into SVD factored form via a series of simple steps based on understanding the geometry underlying the SVD.

Example 1. \[ A = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \]. First note that \( m = 2, n = 1, r = r(A) = 1 \) (obviously),
\( \nu = n - r = 0, \) and \( \mu = m - r = 1 \). This immediately tells us that \( V = V_1 = v_1 = 1 \).

We have

\[
\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \cdot \sqrt{5} \cdot \begin{pmatrix} 1/\sqrt{5} \\ 0 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \cdot \begin{pmatrix} X \\ X \end{pmatrix} \cdot \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1/\sqrt{5} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix} \cdot 1 \]

Note that we exploit the fact that we know the dimensions of the various matrices we have to compute. Here we first filled out \( \Sigma \) before determining the unknown values of \( U_2 = u_2, \) which was later done using the fact that \( u_2 \perp u_1 = U_1. \)
Two Simple SVD Examples – Cont.

Example 2. \( A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \). Note that \( m = 2, n = 2, r = r(A) = 1 \) (obviously), \( \nu = n - r = 1 \), and \( \mu = m - r = 1 \). Unlike the previous example, here we have a nontrivial nullspace.

\[
\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \cdot 2 \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & X \\ \frac{\sqrt{2}}{2} & X \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ X \\ X \end{pmatrix}
\]

Exploiting the facts that \( U_1 = u_1 \perp u_2 = U_2 \) and \( V_1 = v_1 \perp v_2 = V_2 \) we easily determine that

\[
\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}
\]

Note the ± sign ambiguity in the choice of \( U_2 \) and \( V_2 \).