# Lecture 4
# SINGULAR VALUE DECOMPOSITION

Harrison H. Barrett
University of Arizona

Motivation:

Convolution operators can be converted to simple multiplication by Fourier analysis

*Any* linear operator can be converted to simple multiplication by SVD analysis

Thus SVD is a generalization of Fourier analysis; Fourier is a special case of SVD.

## OUTLINE

- Review of lecture 3
  - Meaning of adjoint
  - Importance of being Hermitian
  - How to make a Hermitian operator if nature doesn't give you one.
- Some notation – inner and outer products
- Eigenanalysis of $\mathcal{H}^\dagger \mathcal{H}$ and $\mathcal{H}\mathcal{H}^\dagger$
- Exact (but useful) representations of objects and images
- Singular value decompositions of $\mathcal{H}$ and $\mathcal{H}^\dagger$
- The SVD imaging equation

## REFERENCE

H. H. Barrett and K. J. Myers, Foundations of Image Science (Wiley, 2004)

Chap. 1: Vectors and Operators

Chap. 7: Deterministic Descriptions of Imaging Systems

# Adjoints

The adjoint of a linear operator maps an image vector to *something* in object space (not usually the original object that produced the image).

The adjoint of a general linear operator is defined by

$$(\mathbf{g}_2, \boldsymbol{\mathcal{H}}\mathbf{f}_1)_{\mathbb{V}} = (\boldsymbol{\mathcal{H}}^\dagger \mathbf{g}_2, \mathbf{f}_1)_{\mathbb{U}} \tag{1.39}$$

For a matrix, the adjoint is the complex conjugate of the transpose.

For a CC operator, the forward and adjoint operators are:

$$[\boldsymbol{\mathcal{H}}\mathbf{f}](\mathbf{r}_d) = g(\mathbf{r}_d) = \int_\infty d^q r \; h(\mathbf{r}_d, \mathbf{r})\, f(\mathbf{r})$$

$$\left[\boldsymbol{\mathcal{H}}^\dagger \mathbf{g}\right](\mathbf{r}) = \int_\infty d^s r_d \; h^{(\dagger)}(\mathbf{r}, \mathbf{r}_d)\, g(\mathbf{r}_d) = \int_\infty d^s r_d \, h^*(\mathbf{r}_d, \mathbf{r})\, g(\mathbf{r}_d)$$

So the *kernel* of the adjoint operator is the complex conjugate of the forward kernel with arguments transposed: $h^{(\dagger)}(\mathbf{r}, \mathbf{r}_d) = h^*(\mathbf{r}_d, \mathbf{r})$

# Adjoints – cont.

For a CD operator, the forward and adjoint operators are:

Forward operator – weight continuous function with sensitivity function and integrate:

$$[\boldsymbol{\mathcal{H}}\mathbf{f}]_m = \int_\infty d^q r \; h_m(\mathbf{r}) \, f(\mathbf{r})$$

Adjoint operator – weight discrete vector with complex conjugate of sensitivity function and sum:

$$\left[\boldsymbol{\mathcal{H}}^\dagger \mathbf{g}\right](\mathbf{r}) = \sum_{m=1}^{M} h_m^*(\mathbf{r}) \, g_m$$

## Importance of being Hermitian

Compact Hermitian operators (see Barrett and Myers, Sec. 1.3.3) have a countable set of eigenvectors and eigenvalues:

$$\boldsymbol{\mathcal{A}}\boldsymbol{\psi}_n = \lambda_n \boldsymbol{\psi}_n \,. \tag{1.67}$$

• Eigenvalues are real

• Eigenvectors are orthonormal: $(\boldsymbol{\psi}_n, \boldsymbol{\psi}_m) = \delta_{nm}$

• Eigenvectors form a basis for domain of $\boldsymbol{\mathcal{A}}$, so any vector in the domain can be expressed as superposition of eigenvectors:

$$\mathbf{f} = \sum_{n=1}^{N} \alpha_n \boldsymbol{\psi}_n$$

• Hence action of $\boldsymbol{\mathcal{A}}$ on any function is easy to calculate

$$\boldsymbol{\mathcal{A}}\mathbf{f} = \sum_{n=1}^{N} \lambda_n \alpha_n \boldsymbol{\psi}_n$$

## Construction of Hermitian operators

From any linear operator $\boldsymbol{\mathcal{H}}$, we can construct two Hermitian operators, $\boldsymbol{\mathcal{H}}^\dagger\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger$.

Example – $\boldsymbol{\mathcal{H}}$ a CD operator, $\mathbb{U} \to \mathbb{V}$

$\boldsymbol{\mathcal{H}}^\dagger\boldsymbol{\mathcal{H}}$ is a CC operator mapping $\mathbb{U} \to \mathbb{U}$ with kernel given by

$$[\boldsymbol{\mathcal{H}}^\dagger\boldsymbol{\mathcal{H}}](\mathbf{r}, \mathbf{r}') = \sum_{m=1}^{M} h_m^*(\mathbf{r})\, h_m(\mathbf{r}')$$

$\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger$ is a DD operator ($M \times M$ matrix), mapping $\mathbb{V} \to \mathbb{V}$, with matrix elements given by

$$[\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger]_{mm'} = \int d^2r\, h_m(\mathbf{r})\, h_{m'}^*(\mathbf{r})$$

## Notation: adjoints of vectors and scalar products

An $N$D vector $\mathbf{a}$ can be regarded as an $N$-element column vector or equivalently as an $N \times 1$ matrix. Example for $N = 3$:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

The transpose $\mathbf{a}^t$ and adjoint $\mathbf{a}^\dagger$ can be defined just as for any matrix. The transpose or adjoint of an $N \times 1$ matrix (column vector) is a $1 \times N$ matrix (row vector):

$$\mathbf{a}^t = (a_1, a_2, a_3) \ , \qquad \mathbf{a}^\dagger = (a_1^*, a_2^*, a_3^*)$$

Thus a scalar product can be written as

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\dagger \mathbf{b} = (a_1^*, a_2^*, a_3^*) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \sum_{n=1}^{3} a_n^* b_n$$

## Outer or tensor products

An inner or scalar product, denoted $\mathbf{a}^\dagger\mathbf{b}$ is possible only if $\mathbf{a}$ and $\mathbf{b}$ have the same dimension:

$$\mathbf{a} : N \times 1\,, \qquad \mathbf{b} : N \times 1\,, \qquad \mathbf{a}^\dagger\mathbf{b} : (1 \times N) \times (N \times 1) = 1 \times 1 = \text{scalar}$$

An outer or tensor product, denoted $\mathbf{b}\mathbf{a}^\dagger$, on the other hand, can be defined with any two vectors:

$$\mathbf{a} : N \times 1\,, \qquad \mathbf{b} : M \times 1\,, \qquad \mathbf{b}\mathbf{a}^\dagger : (M \times 1) \times (1 \times N) = M \times N \text{ matrix}$$

## $\boxed{\text{Outer product – example}}$

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \qquad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

$4 \times 1$ $\qquad\qquad$ $3 \times 1$

$$\mathbf{b}\mathbf{a}^\dagger = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \begin{bmatrix} a_1^* & a_2^* & a_3^* \end{bmatrix} = \begin{bmatrix} b_1 a_1^* & b_1 a_2^* & b_1 a_3^* \\ b_2 a_1^* & b_2 a_2^* & b_2 a_3^* \\ b_3 a_1^* & b_3 a_2^* & b_3 a_3^* \\ b_4 a_1^* & b_4 a_2^* & b_4 a_3^* \end{bmatrix}$$

$4 \times 3$

$$[\mathbf{b}\mathbf{a}^\dagger]_{mn} = b_m \, a_n^* \qquad \text{Rank} = 1$$

## Outer products as operators

Any matrix can be regarded as an operator, and $\mathbf{b}\mathbf{a}^\dagger$ is no exception. If $\mathbf{b}\mathbf{a}^\dagger$ is $M \times N$, its operand $\mathbf{c}$ must be $N \times 1$ (i.e., the same size as $\mathbf{a}$), and the result of the operation is

$$\mathbf{b}\mathbf{a}^\dagger\mathbf{c} = \mathbf{b}\left(\mathbf{a}^\dagger\mathbf{c}\right) = \left(\mathbf{a}^\dagger\mathbf{c}\right)\mathbf{b} = \text{scalar} \times \mathbf{b}$$

since $\mathbf{a}^\dagger\mathbf{c}$ is the scalar product of $\mathbf{a}$ and $\mathbf{c}$, which can be placed on either side of the vector $\mathbf{b}$. Thus $\mathbf{b}\mathbf{a}^\dagger\mathbf{c}$ is parallel to $\mathbf{b}$ for all $\mathbf{c}$.

With this interpretation, the concept of outer product can be extended to vectors in *any* two Hilbert spaces. Let $\mathbf{u}$ and $\mathbf{f}$ be arbitrary vectors in $\mathbb{U}$ and $\mathbf{v}$ be an arbitrary vector in $\mathbb{V}$. Then $\mathbf{v}\mathbf{u}^\dagger$ maps $\mathbb{U}$ to $\mathbb{V}$ as follows:

$$\mathbf{v}\mathbf{u}^\dagger\mathbf{f} = \mathbf{v}\left(\mathbf{u}^\dagger\mathbf{f}\right) = \left(\mathbf{u}^\dagger\mathbf{f}\right)\mathbf{v} = \text{scalar} \times \mathbf{v}$$

Since we know how to compute scalar products of functions, this definition works for infinite-dimensional Hilbert spaces (function spaces) as well as finite-dimensional spaces.

## Projection operators and basis vectors

On the last slide we saw that $\mathbf{b}\mathbf{a}^\dagger$ acting on any commensurable vector $\mathbf{c}$ always yields a vector parallel to $\mathbf{b}$; it *projects* $\mathbf{c}$ onto the direction of $\mathbf{b}$.

An important special case is when the set $\{\mathbf{u}_n, n = 1.,,,N\}$ is an orthonormal basis for the space $\mathbb{U}$, so that any vector in that space can be written as

$$\mathbf{f} = \sum_{n=1}^{N} \alpha_n \, \mathbf{u}_n = \sum_{n=1}^{N} (\mathbf{u}_n^\dagger \mathbf{f}) \, \mathbf{u}_n = \sum_{n=1}^{N} \mathbf{u}_n \, (\mathbf{u}_n^\dagger \mathbf{f}) = \sum_{n=1}^{N} \mathbf{u}_n \mathbf{u}_n^\dagger \, \mathbf{f}$$

For orthonormal functions, therefore, $\mathbf{u}_n \mathbf{u}_n^\dagger$ projects its operand onto the vector $\mathbf{u}_n$, and if the set is complete, then

$$\sum_{n=1}^{N} \mathbf{u}_n \mathbf{u}_n^\dagger = \mathbf{I} \, ,$$

where $\mathbf{I}$ is the *identity* or unit operator.

## $\boxed{\text{On to SVD!}}$

SVD begins with eigenalysis of the Hermitian operators $\mathcal{H}^\dagger\mathcal{H}$ and $\mathcal{H}\mathcal{H}^\dagger$. (See Barrett and Myers Sec. 1.5.)

The eigenvalue problem for $\mathcal{H}^\dagger\mathcal{H}$ is:

$$\mathcal{H}^\dagger\mathcal{H}\mathbf{u}_n = \mu_n\,\mathbf{u}_n \qquad (1.111)\,.$$

We know that the set $\{\mathbf{u}_n\}$ is orthonormal and complete in $\mathbb{U}$ and that the eigenvalues are real. Moreover, it is not hard to show that

$$\mu_n \geq 0$$

and hence that $\mathcal{H}^\dagger\mathcal{H}$ is non-negative definite.

If $\mu_n = 0$, then

$$\mathcal{H}^\dagger\mathcal{H}\mathbf{u}_n = 0$$

and $\mathbf{u}_n$ is a null vector of $\mathcal{H}^\dagger\mathcal{H}$.

## $\boxed{\text{Eigenanalysis of } \mathcal{H}\mathcal{H}^\dagger}$

Suppose we have solved the eigenvalue problem for $\mathcal{H}^\dagger\mathcal{H}$:

$$\mathcal{H}^\dagger\mathcal{H}\mathbf{u}_n = \mu_n\,\mathbf{u}_n\,.$$

Operating on both sides with $\mathcal{H}$ and throwing in some parentheses, we see that

$$\mathcal{H}\mathcal{H}^\dagger(\mathcal{H}\mathbf{u}_n) = \mu_n\,(\mathcal{H}\mathbf{u}_n)\,,$$

Thus $(\mathcal{H}\mathbf{u}_n)$ is an eigenvector of $\mathcal{H}\mathcal{H}^\dagger$ with the same eigenvalue $\mu_n$, and we can write

$$\mathcal{H}\mathcal{H}^\dagger\,\mathbf{v}_n = \mu_n\,\mathbf{v}_n\,,$$

where $\mathbf{v}_n$ is any constant times $\mathcal{H}\mathbf{u}_n$. If $\mathbf{u}_n$ is not a null vector, we can define

$$\mathbf{v}_n = \frac{1}{\sqrt{\mu_n}}\mathcal{H}\mathbf{u}_n\,, \qquad (\mu_n \neq 0)\,, \tag{1.116}$$

and then the vectors $\{\mathbf{v}_n\}$ will be properly normalized.

## A convention

The eigenvalues of $\mathcal{H}^\dagger\mathcal{H}$ and $\mathcal{H}\mathcal{H}^\dagger$ are real and non-negative, and it is convenient to order them so that

$$\mu_1 \geq \mu_2 \geq .... \geq \mu_R > 0\,, \tag{1.114}$$

where $R$ is the number of nonzero eigenvalues (counting multiplicity).

$R$ is called the rank of the operator. This definition of rank is consistent with the familiar one for matrices, where $R$ is the number of linearly independent rows or columns.

## $\boxed{\text{Summary so far}}$

Eigenvalue problem for $\mathcal{H}^\dagger \mathcal{H}$

$$\mathcal{H}^\dagger \mathcal{H} \mathbf{u}_n = \mu_n \mathbf{u}_n \,, \qquad \mu_n \geq 0$$

Eigenvectors are orthonormal and complete in $\mathbb{U}$:

$$\mathbf{u}_n^\dagger \mathbf{u}_m = \delta_{nm} \,, \qquad \sum_{n=1}^{N} \mathbf{u}_n \mathbf{u}_n^\dagger = \mathcal{I}_{\mathbb{U}} \,.$$

Eigenvalue problem for $\mathcal{H} \mathcal{H}^\dagger$

$$\mathcal{H} \mathcal{H}^\dagger \mathbf{v}_n = \mu_n \mathbf{v}_n \,.$$

Eigenvectors are orthonormal and complete in $\mathbb{V}$:

$$\mathbf{v}_n^\dagger \mathbf{v}_m = \delta_{nm} \,, \qquad \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^\dagger = \mathcal{I}_{\mathbb{V}} \,.$$

If $\mu_n \neq 0$, then

$$\mathbf{v}_n = \frac{1}{\sqrt{\mu_n}} \mathcal{H} \mathbf{u}_n \,, \qquad (\mu_n \neq 0) \,.$$

# Uses of the SVD basis

Since the eigenvectors of $\mathcal{H}^\dagger\mathcal{H}$ are orthonormal and complete in $\mathbb{U}$, an arbitrary object can be written as

$$\mathbf{f} = \sum_{n=1}^{N} \alpha_n \mathbf{u}_n\,, \qquad \alpha_n = \mathbf{u}_n^\dagger \mathbf{f}\,.$$

The effect of $\mathcal{H}^\dagger\mathcal{H}$ on this vector is easy to compute:

$$\mathcal{H}^\dagger\mathcal{H}\mathbf{f} = \sum_{n=1}^{N} \alpha_n \mathcal{H}^\dagger\mathcal{H}\mathbf{u}_n = \sum_{n=1}^{R} \alpha_n\, \mu_n\, \mathbf{u}_n\,,$$

where the upper limit is now $R$ since $\mu_n = 0$ for $n > R$.

Moreover, the effect of the imaging operator $\mathcal{H}$ is also easy to compute:

$$\mathcal{H}\mathbf{f} = \sum_{n=1}^{N} \alpha_n \mathcal{H}\mathbf{u}_n = \sum_{n=1}^{R} \alpha_n \sqrt{\mu_n}\, \mathbf{v}_n\,.$$

## The SVD imaging equation

The (noise-free) imaging equation for an arbitrary linear operator is

$$\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f} \; .$$

The object and image can be expanded in eigenfunctions of $\boldsymbol{\mathcal{H}}^\dagger\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger$, respectively:

$$\mathbf{f} = \sum_{n=1}^{N} \alpha_n \mathbf{u}_n \, , \qquad \mathbf{g} = \sum_{n=1}^{N} \beta_n \mathbf{v}_n \, .$$

But

$$\boldsymbol{\mathcal{H}}\mathbf{u}_n = \sqrt{\mu_n}\,\mathbf{v}_n \quad \text{so} \quad \mathbf{g} = \sum_{n=1}^{R} \sqrt{\mu_n}\,\alpha_n \mathbf{v}_n$$

Thus the coefficients are simply related by

$$\beta_n = \sqrt{\mu_n}\,\alpha_n$$

## $\boxed{\text{SVD representation of the imaging operator } \boldsymbol{\mathcal{H}}}$

The results above can be summarized by saying that

$$\boldsymbol{\mathcal{H}} = \sum_{n=1}^{R} \sqrt{\mu_n}\, \mathbf{v}_n \mathbf{u}_n^{\dagger}\,. \tag{1.120}$$

Proof:

$$\boldsymbol{\mathcal{H}}\mathbf{f} = \sum_{n=1}^{R} \sqrt{\mu_n}\, \mathbf{v}_n \mathbf{u}_n^{\dagger} \sum_{m=1}^{N} \alpha_m \mathbf{u}_m = \sum_{n=1}^{R} \sum_{m=1}^{N} \alpha_m \sqrt{\mu_n}\, \mathbf{v}_n \mathbf{u}_n^{\dagger} \mathbf{u}_m$$

$$= \sum_{n=1}^{R} \sum_{m=1}^{N} \alpha_m \sqrt{\mu_n}\, \mathbf{v}_n\, \delta_{nm} = \sum_{n=1}^{R} \alpha_n \sqrt{\mu_n}\, \mathbf{v}_n\,, \tag{1.124}$$

which agrees with the previous slide.

## $\boxed{\text{Punchline}}$

Use of the eigenvectors of $\boldsymbol{\mathcal{H}}^\dagger\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger$ as basis vectors allows us to reduce the noise-free imaging equation $\mathrm{g} = \boldsymbol{\mathcal{H}}\mathrm{f}$ to a simple multiplication

$$\beta_n = \sqrt{\mu_n}\,\alpha_n$$

and to represent the imaging operator as

$$\boldsymbol{\mathcal{H}} = \sum_{n=1}^{R} \sqrt{\mu_n}\,\mathbf{v}_n\mathbf{u}_n^\dagger\,.$$

Still to come: pseudoinverses, and inverse problems with noise.