Lecture 6

PSEUDOINVERSES

Harrison H. Barrett
University of Arizona

$\boxed{\text{OUTLINE}}$

- Quick review of SVD of linear operators
- Null space and measurement space
- Projection operators
- Moore's definition of Moore-Penrose pseudoinverse
- Penrose equations, Penrose definition of MP pseudoinverse
- SVD definition of Moore-Penrose pseudoinverse

Key results of SVD: object space

$$\boldsymbol{\mathcal{H}}^{\dagger}\boldsymbol{\mathcal{H}}\mathbf{u}_n = \mu_n\,\mathbf{u}_n \tag{1.111}$$

Eigenvalues are real and non-negative ($\mu_n \geq 0$) and can be ordered as:

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

Set $\{\mathbf{u}_n\}$ is orthonormal and complete in ND space $\mathbb{U}$ (N may be ∞):

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

An arbitrary object can be expanded as:

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

Key results of SVD: image space

$$\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^{\dagger}\,\mathbf{v}_m = \mu_m\,\mathbf{v}_m\,,$$

The eigenvalues are the same as for $\boldsymbol{\mathcal{H}}^{\dagger}\,\boldsymbol{\mathcal{H}}$ and hence the rank is the same.

Set $\{\mathbf{u}_m\}$ is orthonormal and complete in MD space $\mathbb{V}$ (M may be ∞):

$$\mathbf{v}_n^{\dagger}\mathbf{v}_m = \delta_{nm}\,, \qquad \sum_{m=1}^{M} \mathbf{v}_m\,\mathbf{v}_m^{\dagger} = \mathbf{I}_{\mathbb{V}}$$

An arbitrary image can be expanded as:

$$\mathbf{g} = \sum_{m=1}^{M} \beta_m \mathbf{v}_m = \sum_{m=1}^{M} \mathbf{v}_m \mathbf{v}_m^{\dagger}\mathbf{g}\,, \qquad \beta_m = \mathbf{v}_m^{\dagger}\mathbf{g}\,.$$

If $\mu_n \neq 0$, object-space and image-space solutions are related by

$$\mathbf{v}_n = \frac{1}{\sqrt{\mu_n}}\boldsymbol{\mathcal{H}}\mathbf{u}_n\,, \qquad \mathbf{u}_n = \frac{1}{\sqrt{\mu_n}}\boldsymbol{\mathcal{H}}^{\dagger}\mathbf{v}_n$$

Representations of operators

The imaging operator and its adjoint are:

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

and the original Hermitian operators are:

$$\boldsymbol{\mathcal{H}}^\dagger \boldsymbol{\mathcal{H}} = \sum_{n=1}^{R} \mu_n \, \mathbf{u}_n \mathbf{u}_n^\dagger\,, \qquad \boldsymbol{\mathcal{H}} \boldsymbol{\mathcal{H}}^\dagger = \sum_{m=1}^{R} \mu_m \, \mathbf{v}_m \mathbf{v}_m^\dagger$$

The unit operators in the two spaces are

$$\mathbf{I}_{\mathbb{U}} = \sum_{n=1}^{N} \mathbf{u}_n \, \mathbf{u}_n^\dagger\,, \qquad \mathbf{I}_{\mathbb{V}} = \sum_{m=1}^{M} \mathbf{v}_m \, \mathbf{v}_m^\dagger$$

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:

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

and the coefficients are related by

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

Hence

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

Null space and measurement space

Object representation (one more time):

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

Now split the sum into two parts:

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

Note that the second sum is a null vector of $\boldsymbol{\mathcal{H}}$:

$$\boldsymbol{\mathcal{H}} \sum_{n=R+1}^{N} \alpha_n \mathbf{u}_n = \sum_{n=R+1}^{N} \alpha_n \boldsymbol{\mathcal{H}} \mathbf{u}_n = \sum_{n=R+1}^{N} \alpha_n \sqrt{\mu_n} \mathbf{v}_n = 0$$

since $\mu_n = 0$ for $n > R$.

Null space and measurement space – cont.

Thus, for a given operator $\boldsymbol{\mathcal{H}}$, any object can be decomposed as

$$\mathbf{f} = \mathbf{f}_{meas} + \mathbf{f}_{null}\,, \qquad \mathbf{f}_{meas} = \sum_{n=1}^{R} \alpha_n \mathbf{u}_n\,, \qquad \mathbf{f}_{null} = \sum_{n=R+1}^{N} \alpha_n \mathbf{u}_n\,.$$

The set of all measurement vectors of the form $\sum_{n=1}^{R} \alpha_n \mathbf{u}_n$ is called *measurement space* and denoted $\mathbb{U}_{meas}$.

The set of all null vectors of the form $\sum_{n=R+1}^{N} \alpha_n \mathbf{u}_n$ is called *null space* and denoted $\mathbb{U}_{null}$.

Any vector in $\mathbb{U}_{meas}$ is orthogonal to any vector in $\mathbb{U}_{null}$, and these two *orthogonal subspaces* make up the overall object space:

$$\mathbb{U} = \mathbb{U}_{meas} \oplus \mathbb{U}_{null}$$

Consistency space and inconsistency space

An arbitrary vector in image space can be represented as:

$$\mathbf{g} = \sum_{m=1}^{M} \beta_m \mathbf{v}_m \,.$$

Again split the sum into two parts:

$$\mathbf{g} = \sum_{m=1}^{R} \beta_m \mathbf{v}_m + \sum_{m=R+1}^{M} \beta_m \mathbf{v}_m \,.$$

Note that the first sum could have been produced by the operator $\boldsymbol{\mathcal{H}}$ acting on *some* object, i.e., there exists an object $\mathbf{f}_0$ such that

$$\boldsymbol{\mathcal{H}}\mathbf{f}_0 = \sum_{m=1}^{R} \beta_m \mathbf{v}_m \,.$$

The first sum is *consistent* with the operator in this sense.

The second sum, on the other hand, must arise entirely from noise; it is the *inconsistent* part of the data.

Consistency space and inconsistency space – cont.

Thus, for a given operator $\boldsymbol{\mathcal{H}}$, any vector in image space can be decomposed as

$$\mathbf{g} = \mathbf{g}_{con} + \mathbf{g}_{incon}\,, \qquad \mathbf{g}_{con} = \sum_{m=1}^{R} \beta_m \mathbf{v}_m\,, \qquad \mathbf{g}_{con} = \sum_{m=R+1}^{M} \beta_m \mathbf{v}_m\,.$$

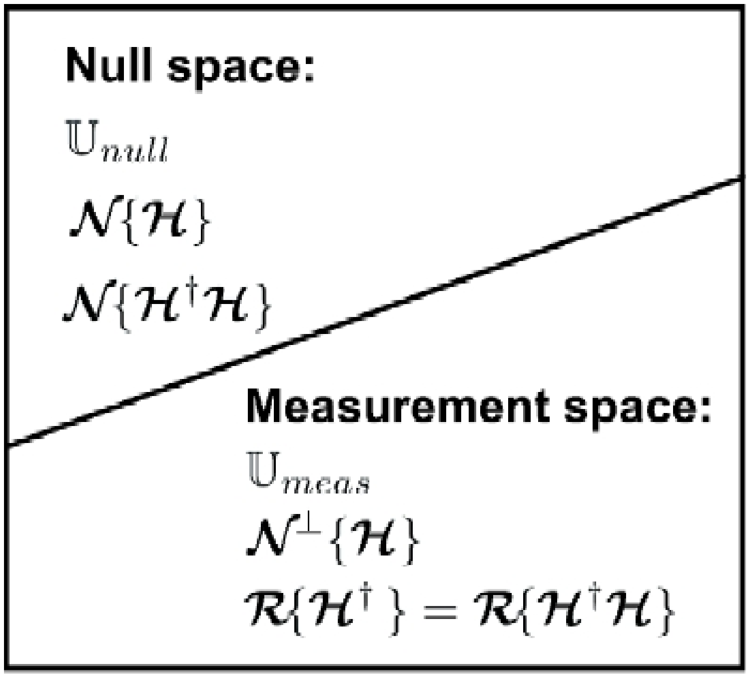
The set of all consistent vectors in image space is called *consistency space* and denoted $\mathbb{V}_{con}$. The set of all inconsistent vectors in image space is called *inconsistency space* and denoted $\mathbb{V}_{incon}$.

Any vector in $\mathbb{V}_{con}$ is orthogonal to any vector in $\mathbb{V}_{incon}$, and these two orthogonal subspaces make up the overall image space:

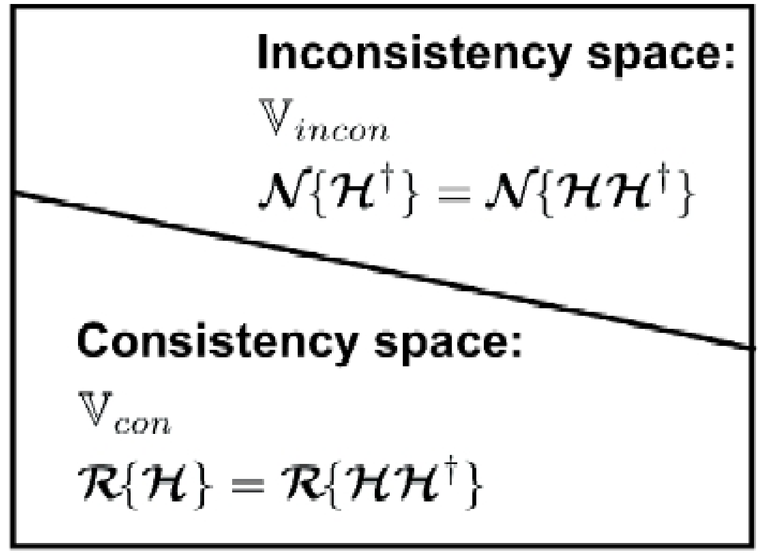
$$\mathbb{V} = \mathbb{V}_{con} \oplus \mathbb{V}_{incon}$$

Exercise: Show that $\boldsymbol{\mathcal{H}}^{\dagger}\mathbf{g}_{incon} = 0$, so inconsistency space is the null space of the adjoint.

Four subspaces



$\mathbb{U}$: Domain of the operator $\mathcal{H}$

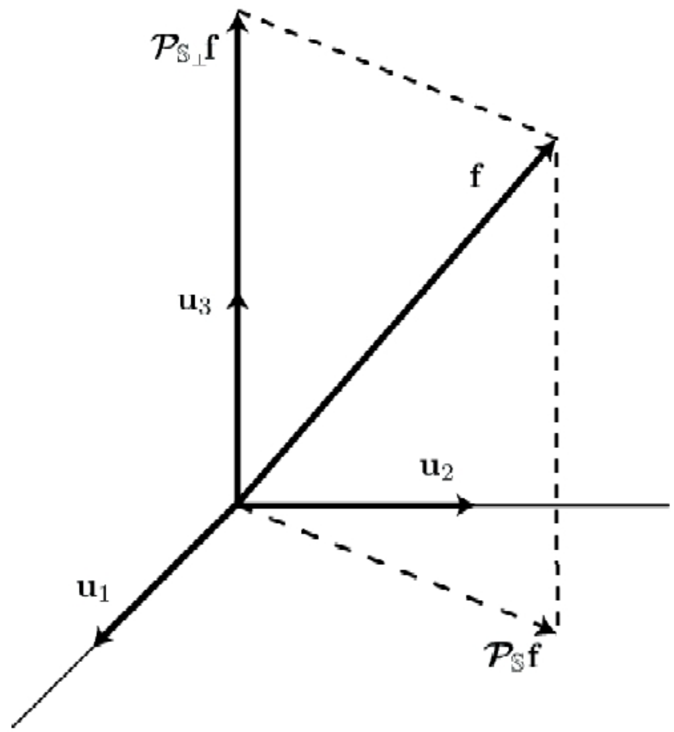


$\mathbb{V}$: Domain of the adjoint operator $\mathcal{H}^\dagger$

Recall $R \leq min(M, N)$. If $R = N$, there is no nontrivial null space. If $R = M$, there is no nontrivial inconsistency space.

Projection operators

A general projection operator $\mathcal{P}$ is Hermitian ($\mathcal{P}^\dagger = \mathcal{P}$) and *idempotent* ($\mathcal{P}^2 = \mathcal{P}$). We are interested here in projections onto subspaces:



Projections onto measurement and null space

Recall:

$$\mathbf{f} = \mathbf{f}_{meas} + \mathbf{f}_{null}\,, \qquad \mathbf{f}_{meas} = \sum_{n=1}^{R} \alpha_n \mathbf{u}_n\,, \qquad \mathbf{f}_{null} = \sum_{n=R+1}^{N} \alpha_n \mathbf{u}_n\,.$$

We can use this decomposition to define two projection operators:

$$\mathbf{f}_{meas} = \boldsymbol{\mathcal{P}}_{meas}\mathbf{f}\,, \qquad \mathbf{f}_{null} = \boldsymbol{\mathcal{P}}_{null}\mathbf{f}\,,$$

where $\boldsymbol{\mathcal{P}}_{meas}$ and $\boldsymbol{\mathcal{P}}_{null}$ are the projectors (or projection operators) onto measurement and null space, respectively.

Caution: A projecor onto a subspace is not the same as a projector onto a convex set and not the same as a projector in tomography.

Exercises: Show that

$$\boldsymbol{\mathcal{P}}_{meas} = \sum_{n=1}^{R} \mathbf{u}_n \mathbf{u}_n^{\dagger}\,, \qquad \boldsymbol{\mathcal{P}}_{null} = \sum_{n=R+1}^{N} \mathbf{u}_n \mathbf{u}_n^{\dagger} = \mathbf{I}_{\mathbb{U}} - \boldsymbol{\mathcal{P}}_{meas}$$

Projections onto consistency and inconsistency space

Recall:

$$\mathbf{g} = \mathbf{g}_{,con} + \mathbf{g}_{\,incon}\,, \qquad \mathbf{g}_{\,con} = \sum_{m=1}^{R} \beta_m \mathbf{v}_m\,, \qquad \mathbf{g}_{\,con} = \sum_{m=R+1}^{M} \beta_m \mathbf{v}_m\,.$$

Define:

$$\mathbf{g}_{\,con} = \boldsymbol{\mathcal{P}}_{\,con}\mathbf{g}\ , \qquad \mathbf{g}_{\,incon} = \boldsymbol{\mathcal{P}}_{\,incon}\mathbf{g}\ ,$$

where $\boldsymbol{\mathcal{P}}_{\,con}$ and $\boldsymbol{\mathcal{P}}_{\,incon}$ are the projectors onto consistency and inconsistency space, respectively.

Exercises: Show that

$$\boldsymbol{\mathcal{P}}_{\,con} = \sum_{m=1}^{R} \mathbf{v}_m \mathbf{v}_m^{\dagger}\,, \qquad \boldsymbol{\mathcal{P}}_{\,incon} = \sum_{m=R+1}^{M} \mathbf{v}_m \mathbf{v}_m^{\dagger} = \mathbf{I}_{\mathbb{V}} - \boldsymbol{\mathcal{P}}_{\,con}$$

Moore's definition of pseudoinverse

Moore (1920) defined an operator $\boldsymbol{\mathcal{H}}^+$ (a matrix in his case) such that:

$$\boldsymbol{\mathcal{P}}_{meas} = \boldsymbol{\mathcal{H}}^+\boldsymbol{\mathcal{H}}\,, \qquad \boldsymbol{\mathcal{P}}_{null} = \mathbf{I}_{\mathbb{U}} - \boldsymbol{\mathcal{H}}^+\boldsymbol{\mathcal{H}},$$

$$\boldsymbol{\mathcal{P}}_{con} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^+\,, \qquad \boldsymbol{\mathcal{P}}_{null} = \mathbf{I}_{\mathbb{V}} - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^+,$$

These four definitions turn out to completely determine $\boldsymbol{\mathcal{H}}^+$.

Exercise: Use what you know about SVD to find an expression for $\boldsymbol{\mathcal{H}}^+$ from the Moore definition.

The Penrose equations

As a graduate student at Cambridge, Roger Penrose (now Sir Roger) proposed that a *generalized inverse* be defined as any operator $\mathcal{H}^{\#}$ that satisfies one or more of the following equations:

$$\text{Penrose Eq. 1: } \mathcal{H}\mathcal{H}^{\#}\mathcal{H} = \mathcal{H} \,. \tag{1.130a}$$

$$\text{Penrose Eq. 2: } \mathcal{H}^{\#}\mathcal{H}\mathcal{H}^{\#} = \mathcal{H}^{\#} \,. \tag{1.130b}$$

$$\text{Penrose Eq. 3: } (\mathcal{H}\mathcal{H}^{\#})^{\dagger} = \mathcal{H}\mathcal{H}^{\#} \,. \tag{1.130c}$$

$$\text{Penrose Eq. 4: } (\mathcal{H}^{\#}\mathcal{H})^{\dagger} = \mathcal{H}^{\#}\mathcal{H} \,. \tag{1.130d}$$

If $\mathcal{H}$ has a true inverse, it satisfies all four of the Penrose equations. An operator that satisfies Penrose Eq. 1 but not the other three is called a 1-inverse of $\mathcal{H}$, one that satisfies Eqs. 1 and 2 is called a (1,2)-inverse, etc. For matrices, a 1-inverse always exists and can be found by Gaussian elimination.

The Penrose definition of the Moore-Penrose pseudoinverse

The Moore-Penrose pseudoinverse, denoted $\boldsymbol{\mathcal{H}}^{+}$, is the generalized inverse that satisfies all four Penrose equations. Thus the Moore-Penrose pseudoinverse is a (1,2,3,4)-inverse and is equal to the true inverse if one exists.

The Moore-Penrose pseudoinverse is also equivalent to the one defined by Moore (otherwise we wouldn't call it that!).

The Moore-Penrose pseudoinverse exists for all operators with finite-dimensional range ($M < \infty$), including CD operators and matrices. It may not exist for some CC operators.

SVD definition

Recall:

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

Can define a pseudoinverse as

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

This operator satisfies both Moore and Penrose definitions. For example

$$\boldsymbol{\mathcal{H}}^+ \boldsymbol{\mathcal{H}} = \left[\sum_{n=1}^{R} \frac{1}{\sqrt{\mu_n}} \mathbf{u}_n \mathbf{v}_n^\dagger \right] \left[\sum_{m=1}^{R} \sqrt{\mu_m}\, \mathbf{v}_m \mathbf{u}_m^\dagger \right]$$

$$= \sum_{n=1}^{R} \sum_{m=1}^{R} \frac{1}{\sqrt{\mu_n}} \sqrt{\mu_m}\, \mathbf{u}_n \mathbf{v}_n^\dagger \mathbf{v}_m \mathbf{u}_m^\dagger = \sum_{n=1}^{R} \mathbf{u}_n \mathbf{u}_n^\dagger = \boldsymbol{\mathcal{P}}_{meas}$$

$\boxed{\text{Still to come}}$

- Iterative computation of the pseudoinverse

- Application to inverse problems

- Effects of noise

- Characterizing the null space

- Characterizing the range

- Consistency conditions