Physical and Mathematical Interpretations of an Adjoint Model with Application to ROMS

Andy Moore
University of Colorado
Boulder

Linearized Systems

- Consider a state-space (the ocean) with state vectors Φ
- Denote $\Phi = (u,v,T,S,\varsigma,N,P,Z,D,\ldots)^T$
- ROMS is just a set of operators:

$$\partial\Phi/\partial t = \mathrm{M}(\Phi) + \mathrm{F}(t)$$

- In general will be nonlinear.
- For many problems it is of considerable theoretical and practical interest to consider perturbations to Φ

- Let $\Phi \to \Phi + \delta\Phi \quad \mathrm{F} \to \mathrm{F}+f$
- In which case:

$$\partial\delta\Phi/\partial t = \left(\partial \mathrm{M}/\partial\Phi\right)\delta\Phi + \mathrm{M}(\delta\Phi) + f(t)$$

- For many problems, it is sufficient to consider small perturbations: $\left|\delta\Phi^2\right| \ll \left|\delta\Phi\right|$ and $\mathrm{M}(\delta\Phi)$ negligible
- The Tangent Linear Equation (TLE):

$$\partial\delta\Phi/\partial t = \left(\partial \mathrm{M}/\partial\Phi\right)\delta\Phi + f(t)$$

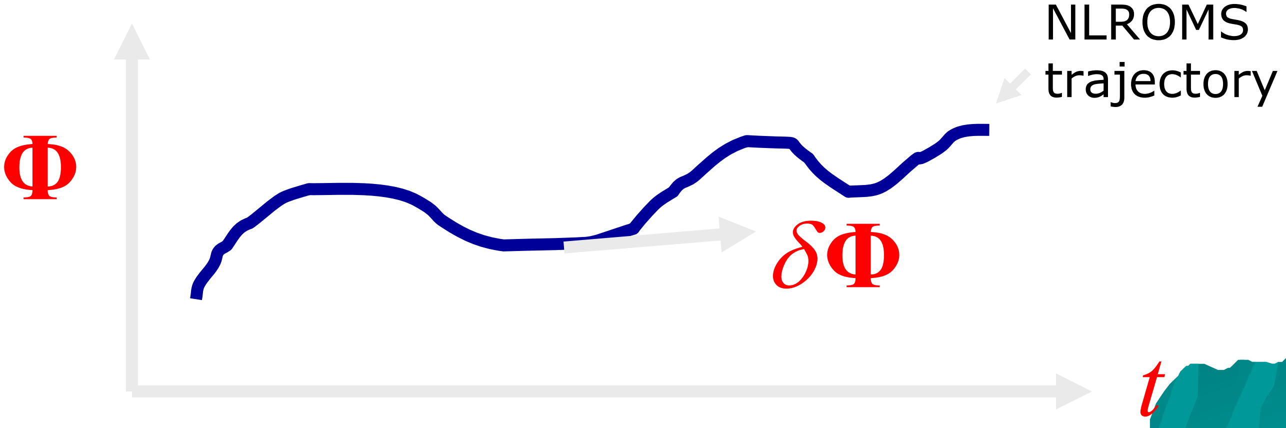
- TLE forms core of many analyses (e.g. normal modes, linear iteration of nonlinear problems (data assimil))

Matrix-Vector Notation

- ROMS solves the primitive equations in discrete form:

$$\partial \mathrm{M}/\partial \Phi \equiv \mathbf{A}(t)$$

$$d\delta\boldsymbol{\Phi}/dt = \mathbf{A}(t)\delta\boldsymbol{\Phi} + \mathbf{f}(t)$$



Important Questions

- Now that we have reduced the linearized ROMS (TLROMS) to a matrix, what would we like to know?
- We should perhaps ask of what value is $\mathbf{A}(t)$ since in reality ROMS (and the real ocean) is nonlinear?

Justification for TLROMS

- All perturbations begin in the linear regime.
- Linear regime often continues to provide useful information long after nonlinearity becomes important.
- Since the action of $\mathrm{M}(\delta\Phi)$ is to merely “scatter” energy, linear regime yields stochastic paradigms.

The Propagator

- It is more convenient to work in terms of the TLROMS propagator:

$$\delta\boldsymbol{\Phi}(t_f) = \mathbf{R}(t_i,t_f)\delta\boldsymbol{\Phi}(t_i)$$

- So, what would we like to know about $\mathbf{R}$?

Dimension

- The ocean is a very large and potentially very complicated place!
- But just how complicated is it?
- What is it’s effective dimension?
- Low dimension described by just a few d.o.f? or high dimension?
- Does dimension depend on where we look?

R is BIG!!!

- R is a monster! $\sim 10^{5-6} \times 10^{5-6}$

R

- Can we reduce R to something more managable?

Enter the Adjoint!

- Eckart-Schmidt-Mirsky theorem: the most efficient representation of a matrix:

$$\mathbf{R} = \sum_{i=1}^{K} \mathbf{u}_i \lambda_i \mathbf{v}_i^T$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the orthonormal singular vectors of $\mathbf{R}$

- Singular Value Decomposition (SVD).

- By definition:

$$\mathbf{R}\mathbf{v}_i = \lambda_i \mathbf{u}_i$$

$$\mathbf{R}^T\mathbf{u}_i = \lambda_i \mathbf{v}_i$$

where λ_i = singular values

$\mathbf{R}^T$ = transpose propagator or adjoint (wrt Euclidean norm)

- Clearly:

$$\mathbf{R}^T\mathbf{R}\mathbf{v}_i = \lambda_i^2 \mathbf{v}_i$$

$$\mathbf{R}\mathbf{R}^T\mathbf{u}_i = \lambda_i^2 \mathbf{u}_i$$

- Consider:

$$\left\langle \mathbf{R}\mathbf{R}^{T} \right\rangle \mathbf{u}_{k} = \nu_{k}\mathbf{u}_{k}$$

- So $\mathbf{u}$ looks like an EOF (i.e. something that we “observe”).

The Question of Dimension

- The dimension of $\mathbf{R}$ is equal to the “range” of $\mathbf{R}$ (i.e. the set of singular vectors with $\lambda_i \neq 0$).
- Dimension=“rank”=maximal # of independent rows and columns of $\mathbf{R}$
- SVD is the most reliable method for determining numerically the rank of a matrix.

Hypothesis

- In the limit $(\Delta x, \Delta y, \Delta z, \Delta t) \to 0$
  $\mathbf{R} \to$ to its continuous counterpart.

- The rank of $\mathbf{R}$ will provide fundamental information about the dimensionality of the real ocean circulation

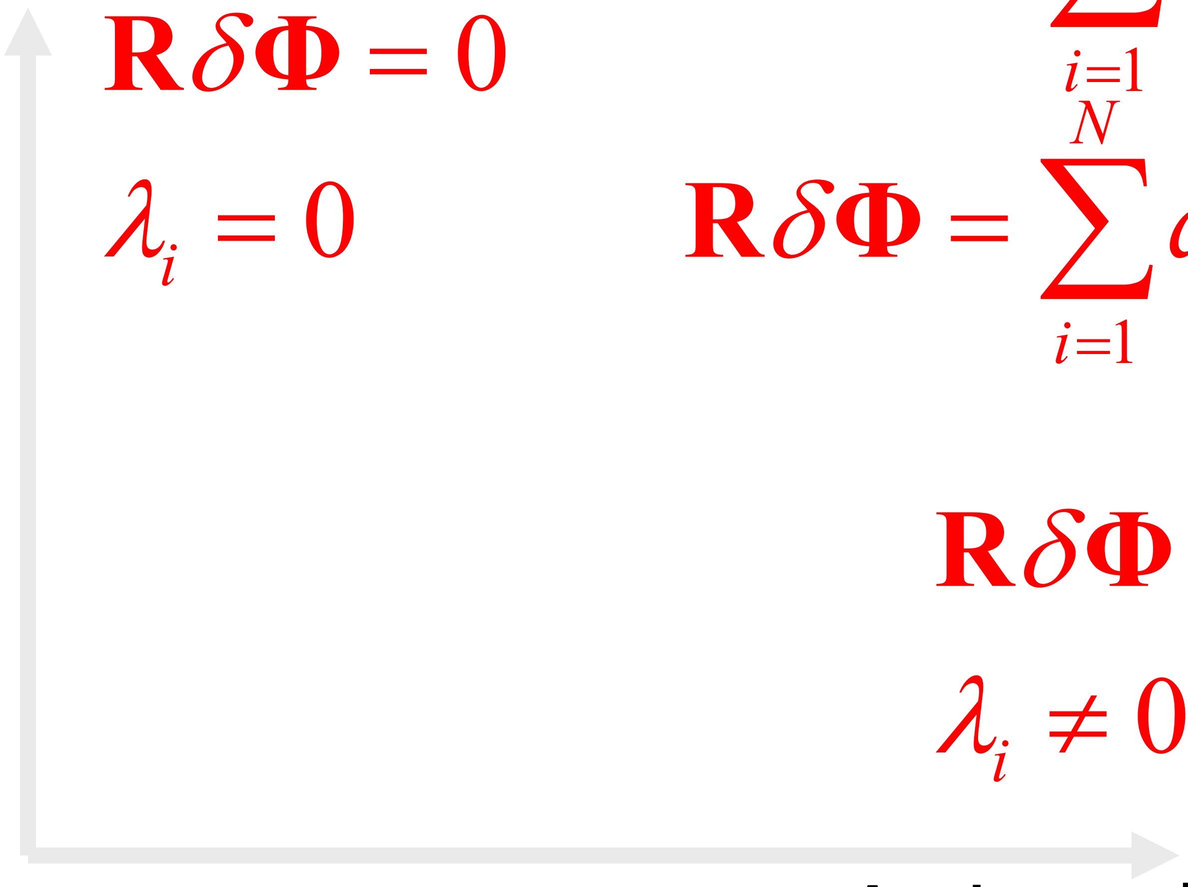
The Active and Null Space

- $\mathbf{R}\mathbf{v}_i = \lambda_i \mathbf{u}_i$ i.e. $\mathbf{R}$ Transforms from v-space to u-space

Initial state

Final/Observed state

- Suppose $\mathbf{R}$ is an (NxN) matrix of rank P (i.e. $\lambda_i \neq 0, i = 1, \ldots, P$
  $\lambda_i = 0, i = P+1, \ldots, N$)
- If $\lambda_i = 0$, $\mathbf{R}\mathbf{v}_i = 0$ (i.e. nothing is observed).
- $\lambda_i = 0$ is the Null Space of $\mathbf{R}$
- $\lambda_i \neq 0$ is the Activated Space of $\mathbf{R}$



Null Space

$$\delta\mathbf{\Phi} = \sum_{i=1}^{N} a_i \mathbf{v}_i$$

$$\mathbf{R}\delta\mathbf{\Phi} = 0$$

$$\lambda_i = 0$$

$$\mathbf{R}\delta\mathbf{\Phi} = \sum_{i=1}^{N} a_i \mathbf{R}\mathbf{v}_i$$

$$\mathbf{R}\delta\mathbf{\Phi} \neq 0$$

$$\lambda_i \neq 0$$

Activated Space

- Recall: $$\mathbf{R}^{T}\mathbf{u}_{i} = \lambda_{i}\mathbf{v}_{i}$$

Observed State

Initial State

- $\mathbf{R}^{T}$ transforms from “observed u” back to “activated v-space”.
- So if we observe “u” the adjoint tells us from whence it came!

  (cf Green’s functions).

Generating Vectors

- Let $\mathbf{A}$ be an (NxM) matrix, where N<M.
- SVD yields two fundamental spaces:

N-space and M-space

$$\underset{(N\times M)(M\times 1)}{\mathbf{A}\mathbf{v}_i} = \lambda_i \underset{(N\times 1)}{\mathbf{u}_i}$$

M-space to N-space

$$\underset{(M\times N)(N\times 1)}{\mathbf{A}^T\mathbf{u}_i} = \lambda_i \underset{(M\times 1)}{\mathbf{v}_i}$$

N-space to M-space

- Consider the underdetermined system

$$\underset{(N\times M)(M\times 1)}{\mathbf{Ax}} = \underset{(N\times 1)}{\mathbf{b}}$$

$\mathbf{A}$, $\mathbf{b}$ given; $\mathbf{x}$ unknown.

- Unique solutions exist if: $\mathbf{x} = \mathbf{A}^T\mathbf{y}$
- Then:

$$\underset{(N\times M)(M\times N)(N\times 1)}{\mathbf{AA}^T\mathbf{y}} = \underset{(N\times 1)}{\mathbf{b}}$$

- $\mathbf{y}$ is called the “generating vector”.
- $\mathbf{x}$ is called the “natural solution”.

- Suppose that $\mathbf{A}$ has only P non-zero singular values: $\lambda_i \neq 0, i = 1, \ldots, P$

- SVD: $$\mathbf{A} = \mathbf{U_p \Lambda_p V_p^T}$$

$$\mathbf{A^T} = \mathbf{V_p \Lambda_p U_p^T}$$

$$\mathbf{x} = \mathbf{A^T y} = \mathbf{V_p (\Lambda_p U_p^T y)} = \mathbf{V_p q}$$

- So $\mathbf{x}$ is ALWAYS in P-space (i.e. “activated space” identified by $\mathbf{A^T}$)

A Familiar Example

- The QG barotropic vorticity equation:

$$\frac{\partial}{\partial t}\left(v_x - u_y\right) + \beta v = 0$$

- Solve for $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$: underdetermined!
- Adjoint vorticity equation yields the generating (stream) function:

$$\mathbf{u} = -\partial\psi/\partial y\,\mathbf{i} + \partial\psi/\partial x\,\mathbf{j}$$

Adjoint Applications

- Clearly the adjoint operator $\mathbf{R}^T$ of ROMS yields information about the subspace or dimensions that are activated by $\mathbf{R}$.
- There are many applications that take advantage of this important property.

Sensitivity Analysis

- Consider a function $J = G(\boldsymbol{\Phi})$
- Clearly $\delta J = \delta\boldsymbol{\Phi}^T\left(\partial G/\partial\boldsymbol{\Phi}\right)$
- But $\delta\boldsymbol{\Phi}(t_f) = \mathbf{R}(t_i,t_f)\delta\boldsymbol{\Phi}(t_i)$
- So $\partial J/\partial\boldsymbol{\Phi} = \mathbf{R}^T(t_f,t_i)\left(\partial G/\partial\boldsymbol{\Phi}\right)$

Sensitivity

- Clearly the action of the adjoint restricts the sensitivity analysis to the subspace activated by $\mathbf{R}^T$ (i.e. to the space occupied by “natural” solutions).

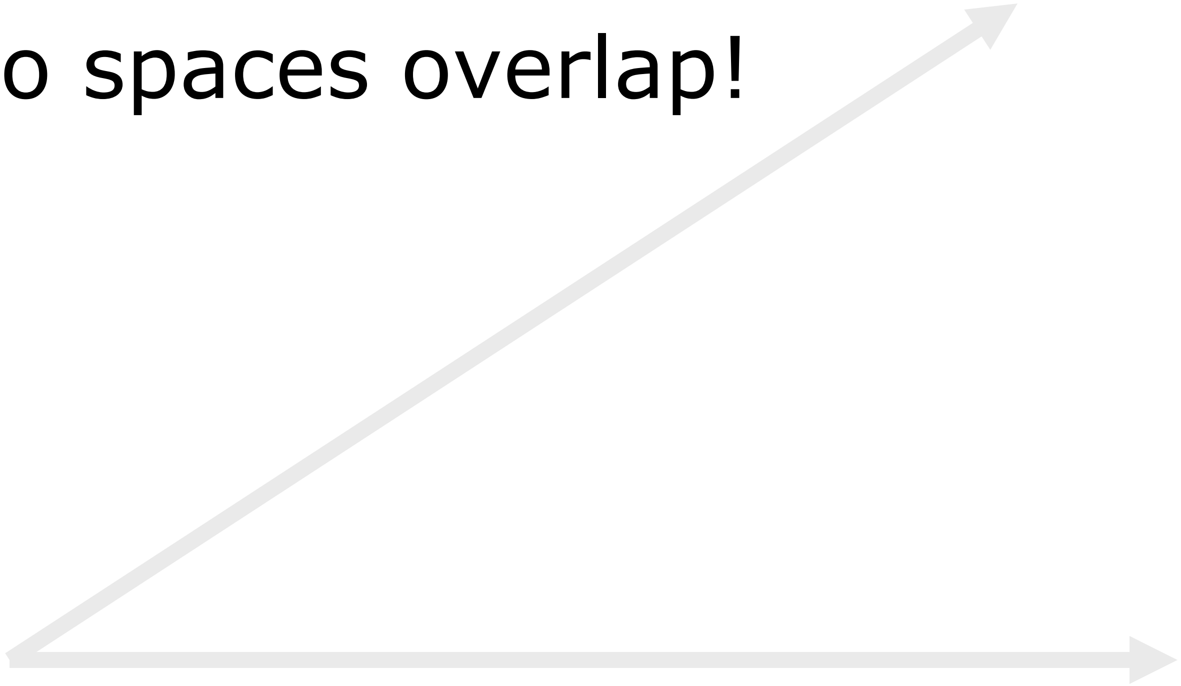
Least-Squares Fitting and Data Assimilation

- If $J = (\mathbf{\Phi} - \mathbf{\Phi}^{obs})^T \mathbf{X} (\mathbf{\Phi} - \mathbf{\Phi}^{obs})$ then the gradient provided by $\mathbf{R}^T$ can be used to find $\mathbf{\Phi}(t_i)$ that minimizes J.
- The is the idea behind 4-dimensional variational data assimilation (4DVAR)
- Clearly the $\mathbf{\Phi}(t_i)$ that minimizes J lies within the active subspace of $\mathbf{R}$

Traditional Eigenmode Analysis

- We are often taught to use the eigenmodes of $\mathbf{R}$ to explore properties and stability of ocean.
- In general, the eigenmodes of $\mathbf{R}$ are NOT orthogonal, meaning each mode has a non-zero projection on other modes.
- What does this do to our notion of active and null space?

The two spaces overlap!

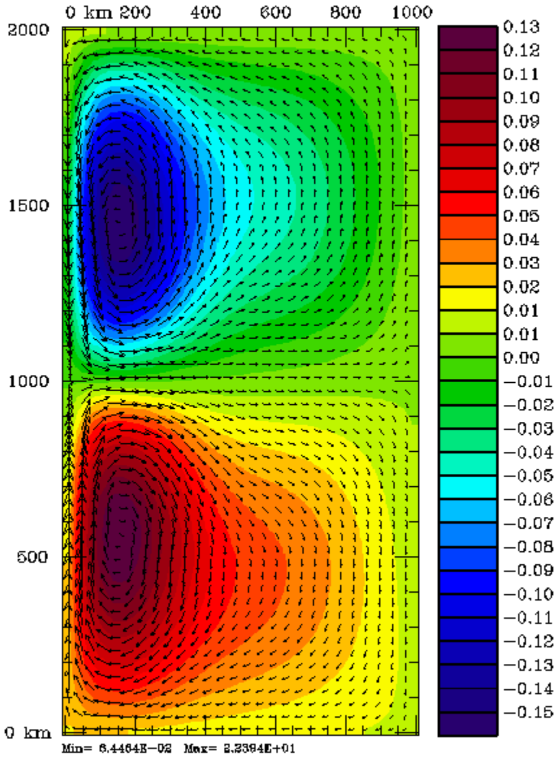


Null Space based on modes of $\mathbf{R}$

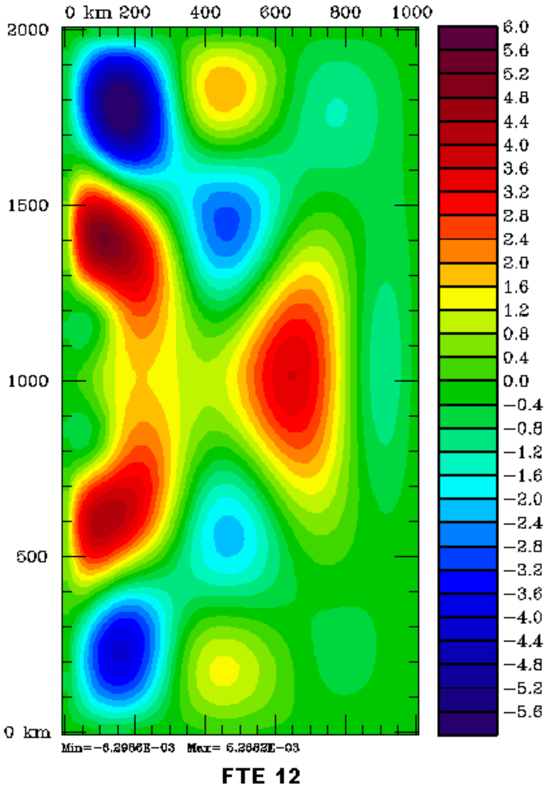
Activated Space based on modes of $\mathbf{R}$

- The amplitude of a particular eigenmode of $\mathbf{R}$ is determined by its projection on the active subspace (i.e. by it’s projection on the corresponding eigenmode of $\mathbf{R}^T$).

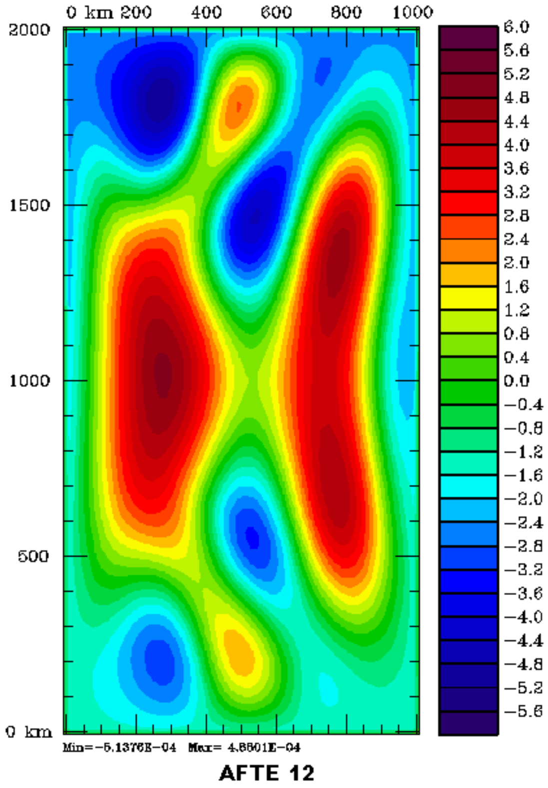
Basin Modes in a Mean Flow



Basic State
Circulation



Eigenmode #6



Adjoint
Eigenmode #6

SVD and Generalized Stability Analysis

- Recall from SVD that: $\mathbf{R}^T\mathbf{R}\mathbf{v}_i = \lambda_i^2\mathbf{v}_i$
- Time evolved SV is $\mathbf{R}\mathbf{v}_i$
- Ratio of final to initial "energy" is:

$$\frac{(\mathbf{R}\mathbf{v}_i)^T(\mathbf{R}\mathbf{v}_i)}{(\mathbf{v}_i^T\mathbf{v})} = \frac{\mathbf{v}_i^T\mathbf{R}^T\mathbf{R}\mathbf{v}_i}{(\mathbf{v}_i^T\mathbf{v})} = \lambda_i^2$$

- So of all perturbations, $\mathbf{v}_1$ is the one that maximizes the growth of "energy" over the time interval $[t_i, t_f]$.

- Consider the forced TL equation:

$$\partial\delta\mathbf{\Phi}\big/\partial t = \left(\partial \mathrm{M}/\partial\mathbf{\Phi}\right)\delta\mathbf{\Phi} + \mathbf{f}(t)$$

- If $\mathbf{f}(t)$ is stochastic in time, more general forms of SVD are of interest.
- Assume unitary forcing: $\left\langle \mathbf{f}(t)\mathbf{f}^T(t)\right\rangle = \mathbf{I}$
- Of particular interest are:

$\mathbf{P} = \int \mathbf{R}\mathbf{R}^T dt$ ⟵ Controllability Grammiam

$\mathbf{Q} = \int \mathbf{R}^T\mathbf{R} dt$ ⟵ Observability Grammiam

- Eigenvectors of $\mathbf{P}$ are the EOFs.
- Eigenvectors of $\mathbf{Q}$ are the Stochastic Optimals.
- Variance: $Var = \mathrm{tr}\left\{\mathbf{P}\right\} = \mathrm{tr}\left\{\mathbf{Q}\right\}$

- Eigenvectors of $\mathbf{P}^{1/2}\mathbf{Q}\mathbf{P}^{1/2}$ are balanced truncation vectors.

- All have considerable practical utility and applications that go far beyond traditional eigenmode analysis!

Covariance Functions and Representers

- The “controllability” Grammiam $\mathbf{P}$ is nothing more than a covariance matrix.
- Note that $\mathbf{b} = \mathbf{P}\mathbf{y} = \left(\int \mathbf{R}\mathbf{R}^T dt\right)\mathbf{y}$ looks a lot like:

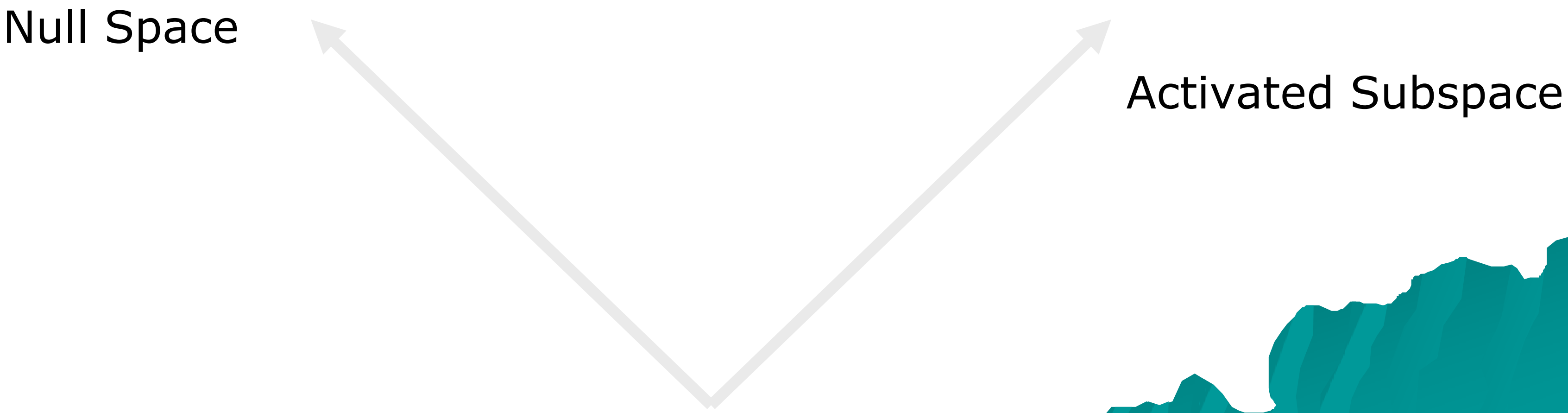
$$\mathbf{A}\mathbf{A}^T\mathbf{y} = \mathbf{b}$$

which yields the “natural solution”.

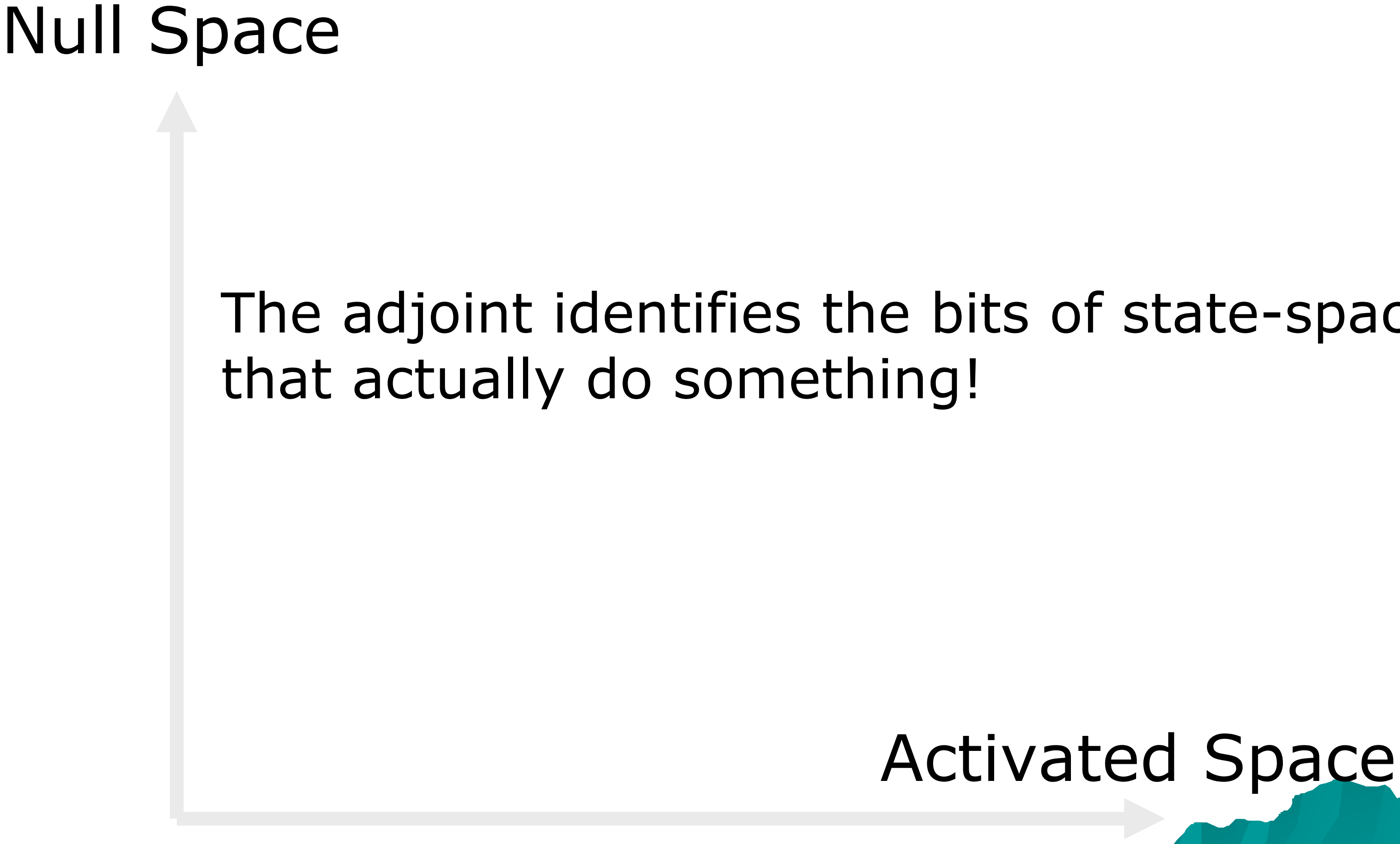
- Operations involving $\mathbf{P}$ yield only natural solutions related to “Representer Functions”.

Norm Dependence

- The adjoint $\mathbf{R}^{\dagger}$ is norm dependent.
- For the Euclidean norm, $\mathbf{R}^{\dagger} = \mathbf{R}^{T}$
- Changing norms is simply equivalent to a rotation and/or change in metric



Summary



Null Space

The adjoint identifies the bits of state-space that actually do something!

Activated Space

The Adjoint of ROMS is a Wonderful Thing!

The Cast of Characters

$M(\Phi)$ - played by NLROMS

$\mathbf{R}$ - played by TLROMS

$\mathbf{R}^T$ - played by ADROMS

Acknowledgements

- Lanczos, C., 1961: Linear Differential Operators, Dover Press.
- Klema, V.C. and A.J. Laub, 1980: The Singular Value Decomposition: Its Computation and Some Applications. IEEE Trans. Automatic Control, AC-25, 164-176.
- Wunsch, C., 1996: The Ocean Circulation Inverse Problem, Cambridge University Press.
- Brogan, W.L., 1991: Modern Control Theory, Prentice Hall.
- Bennett, A.F., 1992: Inverse Methods in Physical Oceanography, Cambridge University Press.
- Bennett, A.F., 2002: Inverse Modeling of the Ocean and Atmosphere, Cambridge University Press.
- Benner, P., P. Mehrmann and D. Sorensen, 2005: Dimension Reduction of Large-Scale Systems. Lecture Notes in Computational Science and Engineering, Vol. 45, Springer-Verlag.
- Trefethen, L.N and M. Embree, 2005: Spectra and Pseudospectra, Princeton University Press.
- Farrell, B.F. and P.J. Ioannou, 1996: Generalized Stability Theory, Part I: Autonomous Operators. J. Atmos. Sci., 53, 2025-2040.
- Farrell, B.F. and P.J. Ioannou, 1996: Generalized Stability Theory, Part II: Nonautonomous Operators. J. Atmos. Sci., 53, 2041-2053.

R is BIG!!!

- R is a monster! $\sim 10^{5-6} \times 10^{5-6}$
- Can we reduce R to something more managable?