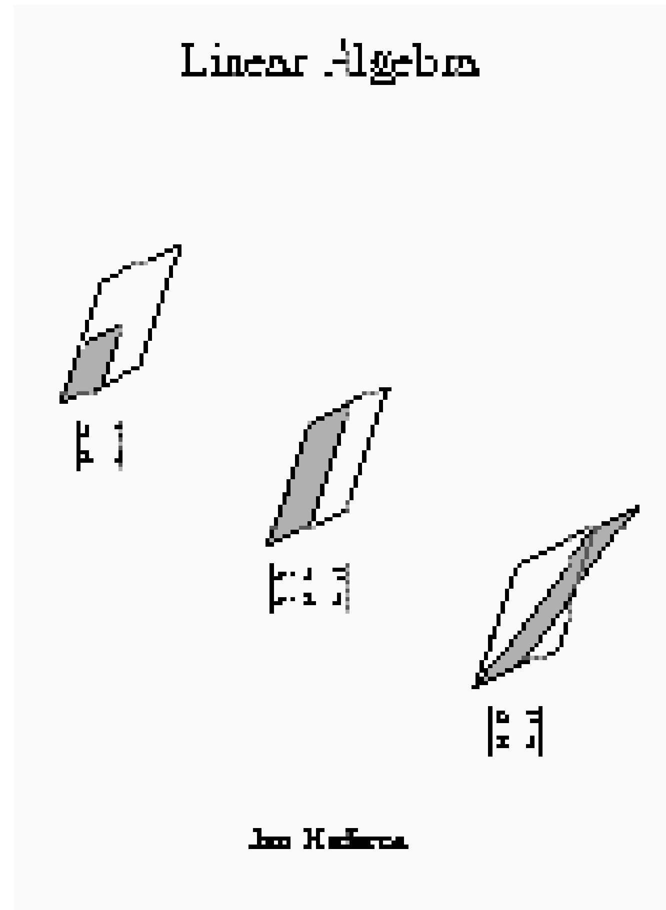
Control Room Accelerator Physics

Day 2

Introduction to Linear Algebra

Outline

1. Introduction
2. Vector spaces
3. Matrices as linear operators
4. Eigenvalues and eigenvectors
5. Diagonalization
6. Singular-value decomposition



Linear Algebra
Introduction

- **Linear algebra** is the branch of mathematics concerned with the study of *vectors*, *vector spaces* (also called *linear spaces*), *linear maps* (also called *linear transformations*, *linear operators*), and systems of linear equations.

- Here we will think of “linear algebra” loosely as matrices
  - Matrices are more “tangible”,
  - They are computation friendly
  - Represent linear maps between finite dimensional vector spaces (e.g., the space of correctors and the space of BPMs)

- Our objective here is to review some basic facts about matrices and establish the notation

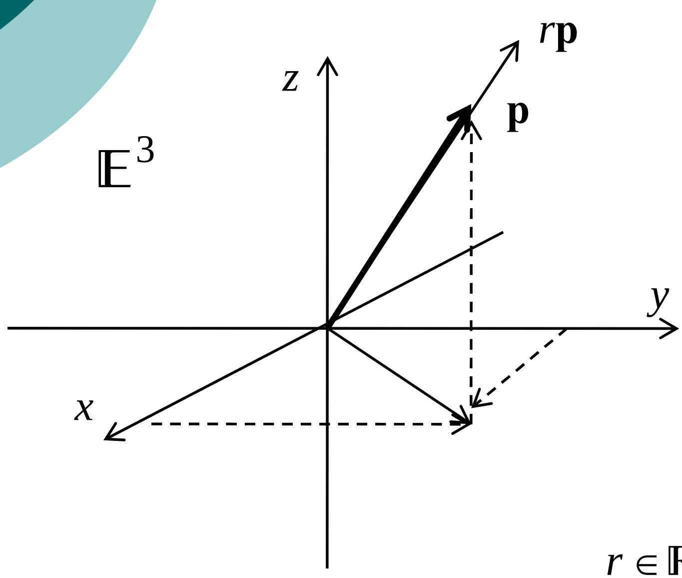
Basic Notation

- $\mathbb{R}$ The set of real numbers (will also use boldface **R**)
- $\mathbb{C}$ The set of complex numbers
  - Numbers of the form $z = \sigma + i\omega$, σ,ω real
- $\mathbb{R}^n$ The Cartesian product of the real line n times or the set of "n-tuples"
  - Vectors of the form $\mathbf{x} = (x_1, \ldots, x_n)$
- $\mathbb{R}^{m\times n}$ The set of $m\times n$ matrices having real number elements
- $GL(n,\mathbb{R})$ The set of elements in $\mathbb{R}^{n\times n}$ having nonzero determinant
  - These are the *invertible* matrices

The Vector Spaces $\mathbf{R}^n$

We can consider the space $\mathbf{R}^n$ as the natural extension of the more familiar 3-vectors in Euclidean 3 space.

$\mathbb{E}^3$



Addition of vectors

$$\mathbf{p} = (p_x, 0, 0) + (0, p_y, 0) + (0, 0, p_z) = (p_x, p_y, p_z)$$

Multiplication by scalars

$$r\mathbf{p} = r(p_x, p_y, p_z) = (rp_x, rp_y, rp_z)$$

$$r \in \mathbb{R}$$

$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

$$\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$$

Vector addition

$$\mathbf{x} + \mathbf{y} = \left(x_1 + y_1, \ldots, x_n + y_n\right)$$

Scalar multiplication

$$r\mathbf{x} = \left(rx_1, \ldots, rx_n\right)$$

Matrices as Linear Operators

Maps Between Vector Spaces $\mathbf{R}^n$

A linear map L has the defining property $L(r\mathbf{x}+s\mathbf{y}) = rL(\mathbf{x}) + sL(\mathbf{y})$ for vectors $\mathbf{x},\mathbf{y}$ and scalars r,s.

- A matrix $\mathbf{A}$ in $\mathbf{R}^{m\times n}$ is, in a natural way, a linear map between the vector spaces $\mathbf{R}^m$ and $\mathbf{R}^n$ under the usual matrix multiplication
- Example:
  - Elements of $\mathbf{R}^m$ are *corrector magnet* strengths
  - Elements of $\mathbf{R}^n$ are *BPM readbacks*
  - Then $\mathbf{A}$ in $\mathbf{R}^{m\times n}$ is the *response matrix*

- We note three important cases:
  - $m > n$ The space $\mathbf{R}^m$ is "bigger" than $\mathbf{R}^n$ (more correctors than BPMs)
  - $m < n$ The space $\mathbf{R}^m$ is "smaller" than $\mathbf{R}^n$ (more BPMs than correctors)
  - $m = n$ The matrix $\mathbf{A}$ is square (equal numbers of correctors and BPMs)

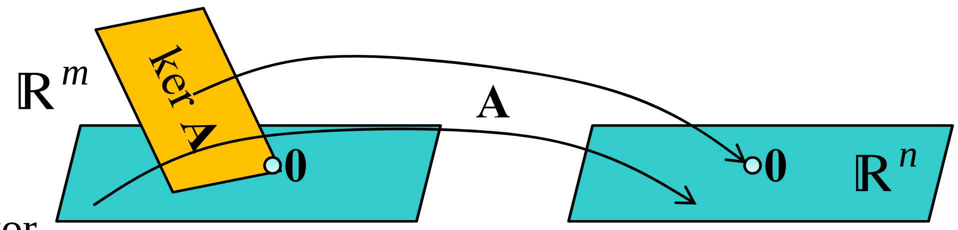
Matrices as Linear Operators

Maps Between Vector Spaces $\mathbf{R}^n$

Case 1: $m>n$

$$\ker \mathbf{A} \equiv \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = 0 \right\} \neq 0$$

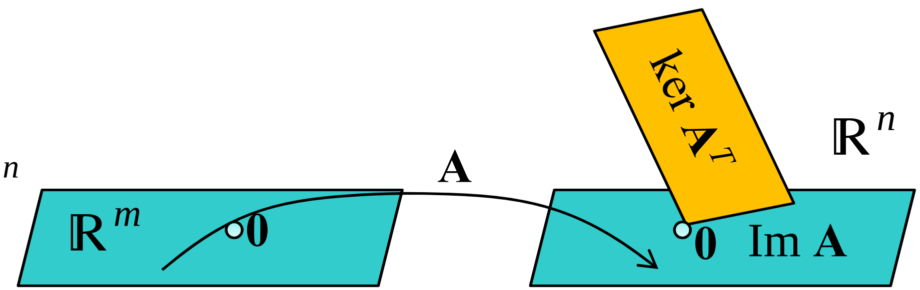
We can get to all beam positions. But, there are many more corrector settings than necessary (and they can "fight" each other).



Case 2: $m<n$

$$\mathrm{Im}\, \mathbf{A} \equiv \left\{ \mathbf{A}\mathbf{x} \in \mathbb{R}^n \mid \text{for all } x \in \mathbb{R}^m \right\} \neq \mathbb{R}^n$$

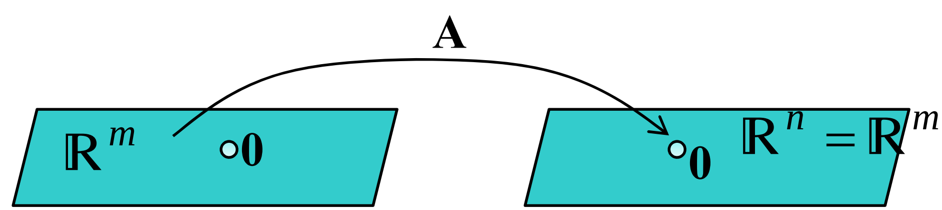
We cannot get to all the beam positions. We don't have enough correctors.



Case 3: $m=n$

$$\ker \mathbf{A} = \mathbf{0}, \quad \mathrm{Im}\, \mathbf{A} = \mathbb{R}^n$$

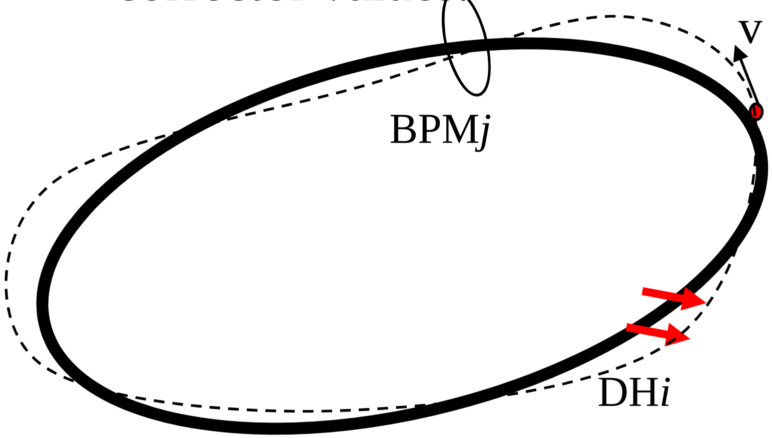
We can steer to all beam positions. Each has a unique corrector setting.



Matrices in Beam Physics

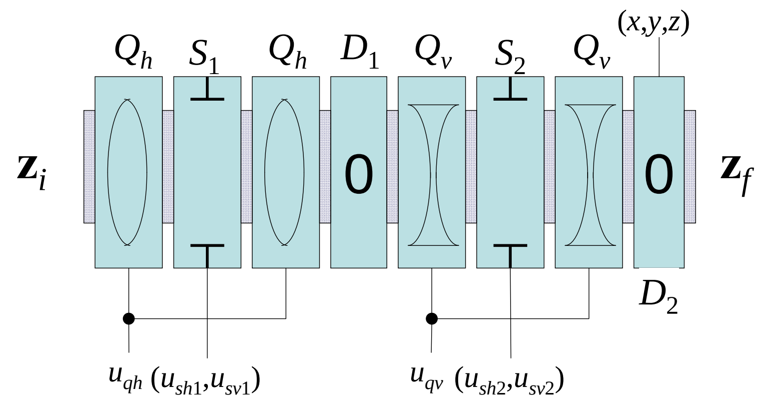
Matrices are Ubiquitous in Beam Physics

Determining the perturbations in the closed orbit of a ring when changing a dipole corrector values.



$$\mathbf{A} = \begin{pmatrix} \dfrac{\partial BPM_1}{\partial DH_1} & \cdots & \dfrac{\partial BPM_1}{\partial DH_m} \\ \vdots & & \vdots \\ \dfrac{\partial BPM_n}{\partial DH_1} & \cdots & \dfrac{\partial BPM_n}{\partial DH_m} \end{pmatrix}$$

Propagating the beam coordinates $\mathbf{z}_f$ at one location of the beamline to another (downstream) location $\mathbf{z}_f$.



$$\mathbf{z}_f = \mathbf{\Phi}(\mathbf{u})\mathbf{z}_i \qquad \mathbf{\Phi} \in \mathbb{R}^{6\times 6}$$

Transfer matrices

Matrix Factorizations

Diagonalization and Singular Value Decomposition

- Because matrices model important aspects of beam physics and control, it is instructive to look at their structure
  - We consider two methods of factoring a matrix, diagonalization and singular-value decomposition
  - Eigenvalues and eigenvectors are intimately related to diagonalization, which we also briefly cover
- Each factorization allows the designer to decouple the effects of the inputs (e.g., corrector strengths) and outputs (e.g., beam positions)
- We focus on response matrices throughout the discussion, however, the XAL online model provides transfer matrices between beamline positions, where this material is also valuable.

Eigenvectors and Eigenvalues
The Natural Modes of A Matrix

- For a *square* matrix $\mathbf{A}$ in $\mathbf{R}^{n\times n}$ we can usually find special vectors $\mathbf{e}$ in $\mathbf{R}^n$ so that

  $$\mathbf{Ae} = \lambda\mathbf{e}$$

  Where $\lambda \neq 0$ is a scalar (either real or complex)

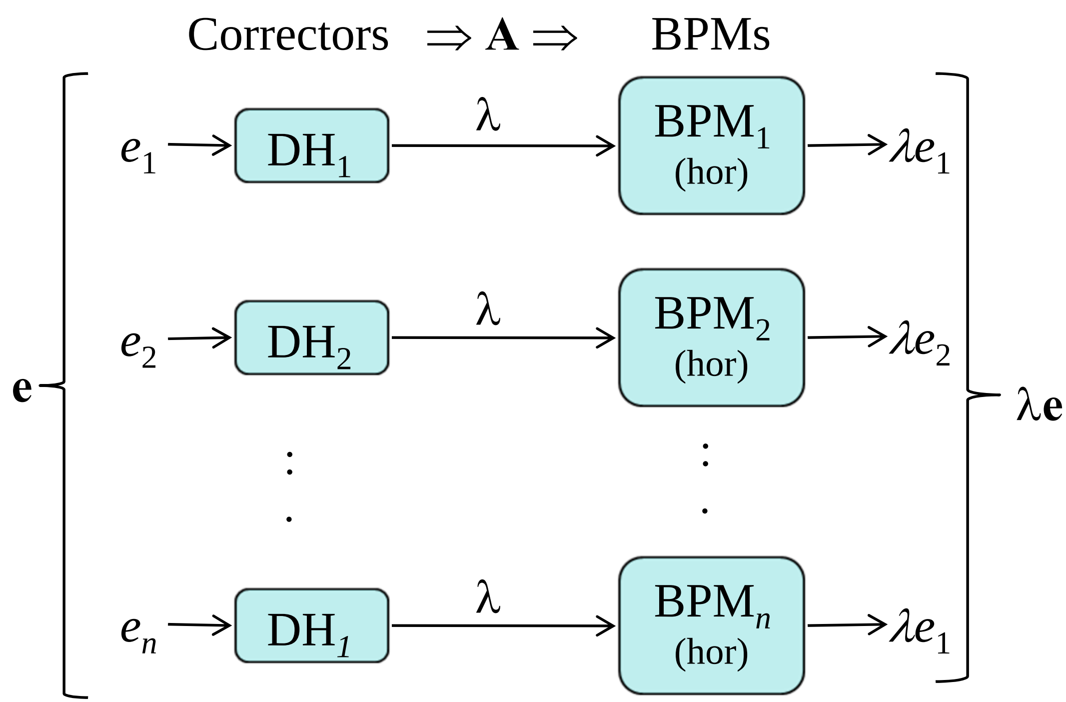
  - Any such vector $\mathbf{e}$ is called an *eigenvector* of $\mathbf{A}$
  - Any such scalar λ is called an *eigenvalue* of $\mathbf{A}$

- $\mathbf{A}$ does not change the direction of $\mathbf{e}$, only the length!
  - $\mathbf{A}$ acts like an amplifier on $\mathbf{e}$ with gain λ
  - The $\{\mathbf{e}\}$ are the natural, uncoupled, modes of $\mathbf{A}$ (as a map)
  - What if we could decompose *all* of $\mathbf{R}^n$ into eigenvectors of $\mathbf{A}$?

Matrix Eigendata
Example: Response Matrix

Say we find an eigenvector **e** for the response matrix **A**

- Drive each dipole corrector DH$_i$ with the value of the eigenvector coordinate e_i
- Then each BPM$_i$ behaves as if it is directly connected to DH$_i$
- The response (beam positions) are simply amplified by the eigenvalue λ



Beamline

This is called a *natural mode* of the system

Matrix Diagonalization
Factoring into Natural Modes

Sometimes a square matrix $\mathbf{A}$ in $\mathbf{R}^{n\times n}$ can be factored as

$$\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$$

Where $\mathbf{T}$ is in $GL(n,\mathbf{R})$ and $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}$ in $\mathbf{R}^{n\times n}$.

- $\mathbf{\Lambda}$ is called the *spectral matrix* (with spectrum $\{\lambda_1,\ldots,\lambda_n\}$)
- $\mathbf{T}$ is called the *modal matrix*
- When this condition is satisfied, i.e., when $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$, we say $\mathbf{A}$ is *diagonalizable*

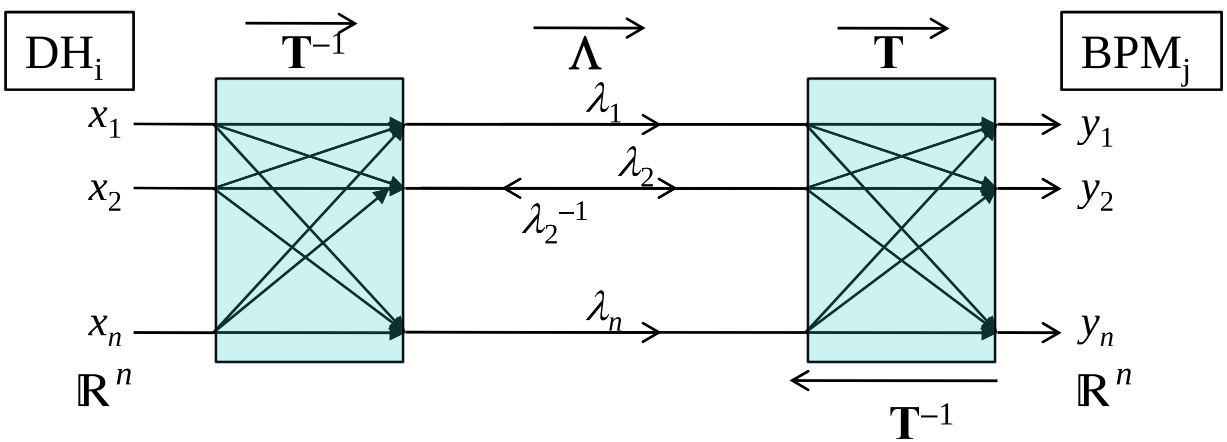
- The matrix $\mathbf{T}$ describes the coupling between the correctors and the BPMs.
- The matrix $\mathbf{\Lambda}$ describes the gains between these natural couplings
- For example….

Matrix Diagonalization

Interpretation

Matrix diagonalization decouples our corrector space ($\mathbf{R}^n$) into the natural modes of the response matrix **A**

- It is possible to inspect the eigenvalues and see which eigenvectors (natural responses) are most sensitive to changes in the correctors
- It is possible to inspect **T** and see which BPMs are most strongly coupled to which correctors



Singular Value Decomposition
Generalizing Diagonalization

Any matrix $\mathbf{A} \in \mathbf{R}^{m\times n}$ may be factored as follows:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$$

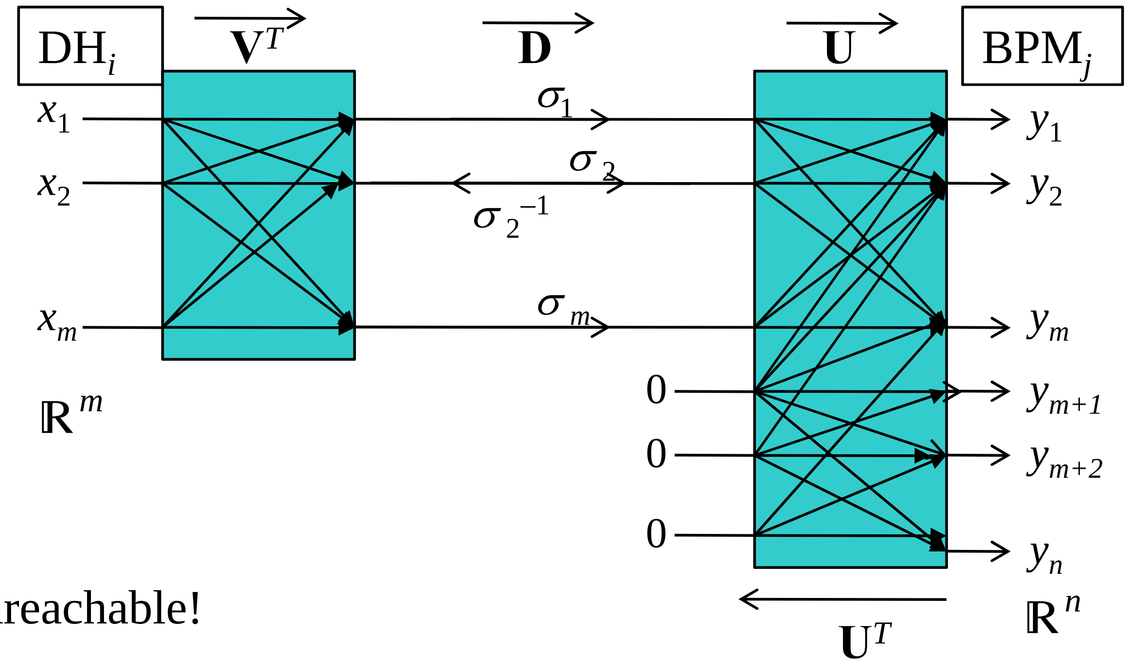
where $\mathbf{U} \in \mathbf{R}^{m\times n}$, $\mathbf{D} = \mathrm{diag}\{\sigma_1,\ldots,\sigma_n\}$, $\mathbf{V} \in \mathbf{R}^{n\times n}$

- The numbers $\{\sigma_1,\ldots,\sigma_n\}$ are the *singular values* of **A**
  - They may be any (complex) number, *including zero*!
  - They are generalizations of eigenvalues for square matrices
- Matrix **U** and **V** have special properties
  - $\mathbf{V}^T\mathbf{V} = \mathbf{I} \in \mathbf{R}^{n\times n}$, that is, it is *orthogonal*
  - $\mathbf{U}^T\mathbf{U} = \mathbf{I} \in \mathbf{R}^{n\times n}$, that is, it is "partially orthogonal"
    - (note $\mathbf{U}^T\mathbf{U} \in \mathbf{R}^{n\times n}$, *not* $\mathbf{R}^{m\times m}$)

Singular Value Decomposition

Generalizing Diagonalization

Returning to the corrector/BPM example with singular-value decomposition



Clearly this situation is analogous to diagonalizable case, however… We must be careful!

- There are outputs which are not connected?!
- Some singular values might be zero!

We have outputs $\{y_i\}$ that are unreachable!

This example was for the case where $m<n$, we have analogous results for $m>n$ (i.e., “dead inputs” and degeneracy)

SVD is an important part of Model-Independent Analysis in high-level beam control

Singular Value Decomposition
Conclusions

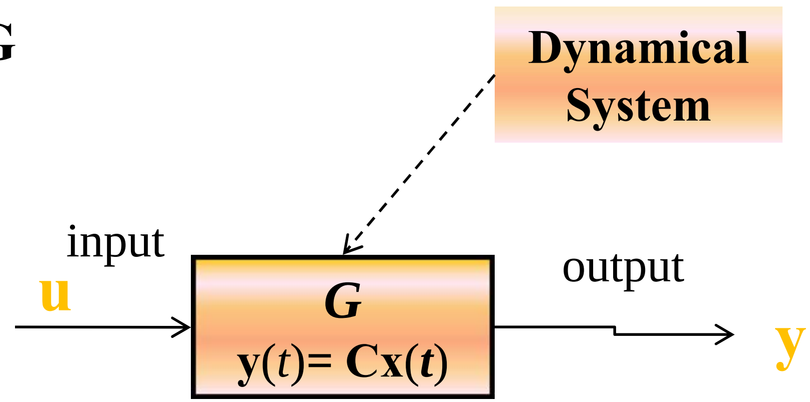
- We can factor *any* matrix $\mathbf{A} \in \mathbf{R}^{m \times n}$ as $\mathbf{A} = \mathbf{UDV}^T$
  - This factorization alone provides an enormous amount of intuition about your system represented by

$$\mathbf{y} = \mathbf{Ax}$$

- Matrix $\mathbf{V}$ tells you which inputs (i.e., the x_i) are viable
  - Correctors that have an effect
- Matrix $\mathbf{U}$ tells you which outputs (i.e., the y_i) are active
  - BPMs that don't respond
- The diagonal matrix $\mathbf{D}$ provides …
  - The gains (i.e., the σ_i) for the system
  - Internal system degeneracy – zero singular value

Linear Systems
A Preview

Linear Systems are again linear maps **G** between vector spaces, but here the vector spaces are dynamic; that is, they contain functions.



- The internal dynamics of linear systems can usually be described by matrix-vector differential equations of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) \qquad \mathbf{x}(t) \in \mathbb{R}^{n}, \;\; \mathbf{A} \in \mathbb{R}^{n\times n}$$

- This may look challenging but we can use all that we have learned here to "disassemble" the above

If $\mathbf{A} = \mathbf{T}\Lambda\mathbf{T}^{-1}$ then $\boldsymbol{\xi} = \mathbf{T}^{-1}$

and

$$\dot{\boldsymbol{\xi}} = \boldsymbol{\Lambda}\boldsymbol{\xi} \quad \Rightarrow \quad \dot{\xi}_i = \lambda_i \xi_i$$

Matrix Exponential

A Preview

- Functions of (square) matrices are common in analysis
  - For example, $\sin(\mathbf{A})$, $\log(\mathbf{A})$, $\exp(\mathbf{A})$, for $\mathbf{A} \in \mathbf{R}^{n\times n}$
  - These functions may seem strange, but they are well-defined by the Taylor series for the function (matrix powers are well-defined)

- Of particular importance for us is the matrix exponential $e^{t\mathbf{A}}$
  - The scalar t representing time
  - This function occurs frequently in linear (dynamical) systems

- Again, say $\mathbf{A}$ is diagonalizable where $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$
  - We shall see that $e^{t\mathbf{A}} = \mathbf{T}e^{t\Lambda}\mathbf{T}^{-1}$
  - The exponential $e^{t\Lambda}$ is easy to compute
  - $e^{t\Lambda} = \mathrm{diag}\{e^{t\lambda 1},\ldots, e^{t\lambda n}\}$

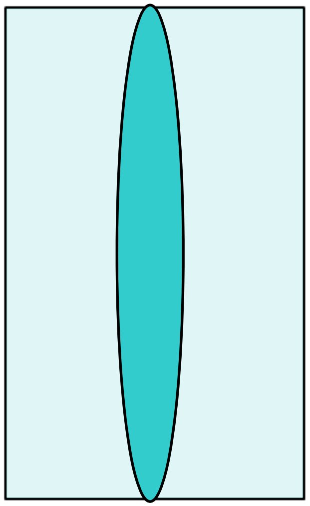
Linear Beam Optics

A Preview

Linear systems and the matrix exponential play a crucial part in linear beam optics and, consequently, the XAL online model

- In linear beam optics beamline elements are modeled by matrices $\mathbf{\Phi}$
- These matrices are formed from the exponential of another matrix $\mathbf{G}$
- The matrix $\mathbf{G}$ represents the equations of motion
- The XAL online model is based upon these ideas

Focusing Quadrupole n



$$\mathbf{z}(s) = \mathbf{\Phi}_n(s)\mathbf{z}_0$$

$$\mathbf{\Phi}_n(s) \equiv e^{s\mathbf{G}_n}$$

$$\mathbf{z}'(s) = \mathbf{G}_n(s)\mathbf{z}(s)$$

Linear Algebra
Summary

- Matrices can be treated as linear operators between finite dimensional vector spaces, in particular, the spaces $\mathbf{R}^n$
- A square matrix $\mathbf{A}$ usually has eigenvalues and eigenvectors that characterize the action of $\mathbf{A}$ upon vector space $\mathbf{R}^n$
- If a matrix $\mathbf{A}$ can be diagonalized as $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$ then its action can be completely decoupled
- *Any* matrix $\mathbf{A} \in \mathbf{R}^{m \times n}$ may be factored according to $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ where $\mathbf{D}$ is the matrix of singular values
  - This factorization, although not as straightforward, also characterizes the action of $\mathbf{A}$ upon $\mathbf{R}^n$ (and domain $\mathbf{R}^m$)

Supplementary Material

- More details on Linear Algebra

Matrix Diagonalization
Factoring into Natural Modes

Sometimes a square matrix $\mathbf{A}$ in $\mathbf{R}^{n\times n}$ can be factored as

$$\mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1}$$

Where $\mathbf{T}$ is in $GL(n,\mathbf{R})$ and $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1,\ldots,\lambda_n\}$ in $\mathbf{R}^{n\times n}$.

- $\boldsymbol{\Lambda}$ is called the *spectral matrix* (with spectrum $\{\lambda_1,\ldots,\lambda_n\}$)
- $\mathbf{T}$ is called the *modal matrix*
- When this condition is satisfied, i.e., when $\mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1}$, we say $\mathbf{A}$ is *diagonalizable*

Fact: If a square matrix $\mathbf{A}$ in $\mathbf{R}^{n\times n}$ has n unique eigenvalues $\{\lambda_1,\ldots,\lambda_n\}$ then it can be factored as above

- In fact, a suitable $\mathbf{T}$ can be formed by augmenting all the eigenvectors as columns. Specifically, $\mathbf{T} = \left(\mathbf{e}_1 \middle| \;\mathbf{e}_2\middle| \;\cdots\; \middle|\mathbf{e}_n\right)$

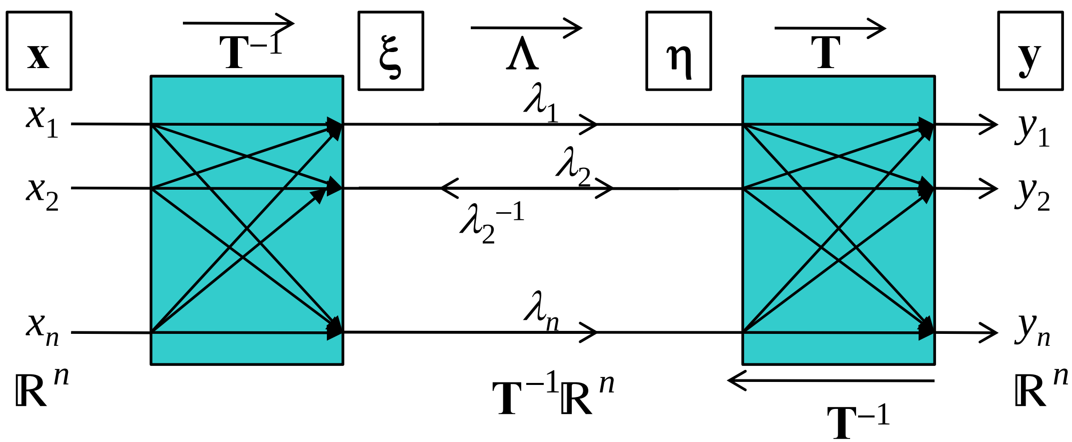
Matrix Diagonalization
Interpretation

Matrix diagonalization decouples $\mathbf{R}^n$ into the natural modes of $\mathbf{A}$

- Instead of working in $\mathbf{R}^n$, we work in the space $\mathbf{T}^{-1}\mathbf{R}^n$!

- To make this less abstract consider equation $\mathbf{y} = \mathbf{A}\mathbf{x}$ and, for example,
  - Think of $\mathbf{A}$ as a multiple-input, multiple-output, coupled amplifier.
  - Instead of using parameters $\mathbf{x}$ and $\mathbf{y}$, use $\xi \equiv \mathbf{T}^{-1}\mathbf{x}$ and $\eta \equiv \mathbf{T}^{-1}\mathbf{y}$
  - Everything decouples as $\xi_i = \lambda_i \eta_i$ (transform back when you're done)

$$\mathbf{y} = \mathbf{T}\Lambda\mathbf{T}^{-1}\mathbf{x}$$
$$\mathbf{T}^{-1}\mathbf{y} = \Lambda\mathbf{T}^{-1}\mathbf{x}$$
$$\boldsymbol{\eta} = \Lambda\boldsymbol{\xi}$$



Matrix Diagonalization

Special Case: Symmetric, Positive Definite Matrix

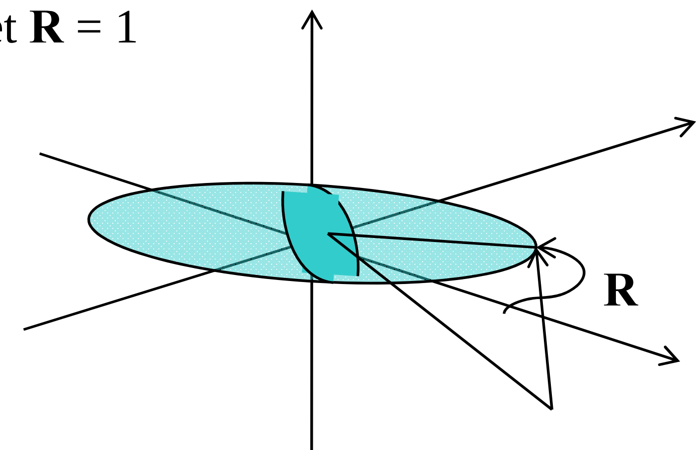
Fact: A positive-definite ($\lambda_i > 0$, for each i), symmetric ($\mathbf{A} = \mathbf{A}^T$), square matrix $\mathbf{A}$ in $\mathbf{R}^{n \times n}$ can always be diagonalized as

$$\mathbf{A} = \mathbf{R}\Lambda\mathbf{R}^T$$

Where $\mathbf{R}$ is in the special orthogonal group $SO(n)$ in $\mathbf{R}^{n \times n}$.

For any element $\mathbf{R}$ of $SO(n)$ in $\mathbf{R}^{n \times n}$

- $\mathbf{R}\mathbf{R}^{\mathrm{T}} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix
- From the above, $\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}$, e.g., just like a rotation in 3-space
- det $\mathbf{R} = 1$



$\mathbf{A}$ appears as a hyper-ellipsoid with semi-axes $\{\lambda_1,\ldots,\lambda_n\}$ rotated by a (generalized) angle $\mathbf{R}$ in hyper-Euclidean n space

Singular Value Decomposition
Generalizing Diagonalization

Any matrix $\mathbf{A} \in \mathbf{R}^{m\times n}$ may be factored as follows:

$$\mathbf{A} = \mathbf{UDV}^T$$

Note the similarity with the special case when **A** is symmetric and positive-definite.

where $\mathbf{U} \in \mathbf{R}^{m\times n}$, $\mathrm{D} = \mathrm{diag}\{\sigma_1,\ldots,\sigma_n\}$, $\mathbf{V} \in SO(n) \in \mathbf{R}^{n\times n}$

- The numbers $\{\sigma_1,\ldots,\sigma_n\}$ are the *singular values* of **A**
  - They may be any (complex) number, *including zero*!
  - They are generalizations of eigenvalues for square matrices
- Matrix **U** has the special property that it is "partially orthogonal"
  - $\mathbf{U}^T\mathbf{U} = \mathbf{I} \in \mathbf{R}^{n\times n}$
  - Note that $\mathbf{V}^T\mathbf{V} = \mathbf{I} \in \mathbf{R}^{n\times n}$ because $\mathbf{V} \in SO(n)$

Singular Value Decomposition
Generalizing Diagonalization

- The columns of $\mathbf{V}$ are called the *right singular vectors* of $\mathbf{A}$
- The columns of $\mathbf{U}$ are called the *left singular vectors* of $\mathbf{A}$

Again consider the matrix-vector equation

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{x}$$

and perform the substitutions

- $\xi = \mathbf{V}^T\mathbf{x}$
- $\eta = \mathbf{U}^T\mathbf{y}$

We then have the (almost) equivalent equation

$$\eta = \mathbf{D}\xi$$

Like the case of a diagonalizable $\mathbf{A}$, this equation is completely decoupled $\Rightarrow$ $\eta_i = \sigma_i \xi_i$