2 Linear Algebra

notation: dot vs transpose

Linear algebra can be thought of as the streelamp of mathematical modeling. In the areas that it illuminates it provides great clarity, but it is also tempting to look there for things that really lie elsewhere. Linear algebra will have a central but uneasy role throughout this book: the assumption of linearity will lead to results that are convenient and understandable, but not applicable to the most challenging problems. Conversely, familiarity with linear algebra can breed a kind of contempt that neglects its depth; this chapter is included rather than assumed as background because of the more advanced results that will be needed elsewhere in the book.

A function $f(x_1, x_2, \ldots, x_N)$ is said to be *linear* if rescaling and offsetting an argument does the same to the result:

$$\begin{aligned} f(x_1, x_2, \ldots, \alpha x_i + \beta, \ldots, x_N) &= \alpha\ f(x_1, x_2, \ldots, x_i, \ldots, x_N) \\ &\quad + \ f(x_1, x_2, \ldots, \beta, \ldots, x_N) \quad . \end{aligned} \tag{2.1}$$

Writing out all of those elements quickly becomes a nuisance, so the first thing that we'll do is group numbers into vectors and matrices. These will be used to solve systems of algebraic equations (and recognize when and why they can't be solved), and then as arguments of nonlinear functions. This chapter will consider static relationships among these variables, which will be extended in the next chapter to dynamic relationships in linear differential and difference equations.

2.1 NUMBERS, VECTORS, AND MATRICES

Modeling usually begins with a (frequently implicit) choice of a way to represent the quantities of interest. This small step can have large consequences, because a bad representation can make an easy problem appear difficult, (and *vice versa*). The question of representation will recur throughout this book; we start here with the simplest question of all, the kind of numbers to use.

Deceptively simple questions about numbers can have remarkably deep answers Wiles naturally counted by integers. While these are usually written

The very first decision to be made in modeling is what kind of numbers to use. If the quantities in a problem are naturally discrete, such as the coordinates of a point on a lattice, or the number of observations in a histogram, use *integers*. These will generally be written in *decimal* notation using powers of ten (i.e., *base 10*), such as

$$249 = 2 \times 10^2 + 4 \times 10^1 + 9 \times 10^0 \quad , \tag{2.2}$$

although when computers are involved it can be more convenient to use

$$11111001 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 \tag{2.3}$$

$$\text{F9} = \text{F}(15) \times 16^1 + 9 \times 16^0 \tag{2.4}$$

base 2 base 16: 0-9A-F
binary
fractions
hexadecimal
rational irrational
[Hardy & Wright, 1998]
complex conjugate ref quaternion chapter *magnitude*
A *vector* $\vec{v}$ is a set of N elements

$$\vec{v} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \quad . \tag{2.5}$$

These might be, for example, the coordinates of a point in an N-dimensional space, or the set of grades of a group of students taking a course in mathematical modeling.
dot product
notation: dot vs transpose
norm measure distances in other spaces angle

$$\begin{aligned} \vec{u} \cdot \vec{v} &= (u_1, \ldots, u_N) \cdot (|\vec{v}|, 0, \ldots, 0) \\ &= u_1 |\vec{v}| \\ &= |\vec{u}||\vec{v}| \cos\theta \end{aligned} \tag{2.6}$$

The *norm* or *magnitude* of a vector $|\vec{v}|$ can be defined as

$$\begin{aligned} |\vec{v}|^2 &= \sum_{n=1}^{N} |v_n|^2 \\ &= \sum_{n=1}^{N} v_n^* v_n \\ &= \sum_{n=1}^{N} v_n^2 \quad \text{(for real variables)} \end{aligned} \tag{2.7}$$

(this is called the L^2 norm because it uses the second power of the components). The *dot product* of two vectors is defined by

$$\vec{u} \cdot \vec{v} = \sum_{n=1}^{N} u_n v_n \tag{2.8}$$

unit vector

A set of vectors $\{\vec{v}_i\}$ is *orthogonal* if

$$\vec{v}_i \cdot \vec{v}_j = 0 \ (i \neq j) \tag{2.9}$$

and *orthonormal* if

$$\vec{v}_i \cdot \vec{v}_j = \delta_{ij} \tag{2.10}$$

(where δ_{ij} is the *Kroenecker delta* which equals 1 if $i = j$ and is 0 otherwise). If all of the columns of a matrix are orthonormal then

$$\mathbf{A}^T\mathbf{A} = \mathbf{I} \equiv \begin{bmatrix} 1 & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & 1 \end{bmatrix} \quad , \tag{2.11}$$

where **I** is the *identity matrix*. A matrix with this property is said to be *orthogonal* (although its columns or rows are really orthonormal).

linear independence

complex norm?

Given a linearly-independent set of vectors $\{u_i\}$, it's possible to construct an orthonormal set $\{v_i\}$ by *Gram-Schmidt orthogonalization*. The first step is to normalize one of the vectors

$$\vec{v}_1 = \frac{\vec{u}_1}{|\vec{u}_1|} \tag{2.12}$$

so that

$$\vec{v}_1 \cdot \vec{v}_1 = \frac{\vec{v}_1 \cdot \vec{v}_1}{|\vec{u}_1|^2} = \frac{|\vec{u}_1|^2}{|\vec{u}_1|^2} = 1 \quad . \tag{2.13}$$

Then this component is subtracted off of a second vector,

$$\vec{v}_2 = \frac{\vec{u}_2 - (\vec{u}_2 \cdot \vec{v}_1)\,\vec{v}_1}{|\vec{u}_2 - (\vec{u}_2 \cdot \vec{v}_1)\,\vec{v}_1|} \quad , \tag{2.14}$$

choosing the magnitude so that

$$\vec{v}_2 \cdot \vec{v}_1 = \frac{\vec{u}_2 \cdot \vec{v}_1 - (\vec{u}_2 \cdot \vec{v}_1)\,\vec{v}_1 \cdot \vec{v}_1}{|\vec{u}_2 - (\vec{u}_2 \cdot \vec{v}_1)\,\vec{v}_1|} = 0 \quad , \tag{2.15}$$

and dividing by the magnitude to normalize $\vec{v}_2 \cdot \vec{v}_2 = 1$. Next, both components are subtracted from a third vector,

$$\vec{v}_3 = \frac{\vec{u}_3 - (\vec{u}_3 \cdot \vec{v}_1)\,\vec{v}_1 - (\vec{u}_3 \cdot \vec{v}_2)\,\vec{v}_2}{|\vec{u}_3 - (\vec{u}_3 \cdot \vec{v}_1)\,\vec{v}_1 - (\vec{u}_3 \cdot \vec{v}_2)\,\vec{v}_2|} \quad ‘ \tag{2.16}$$

with magnitudes chosed so that $\vec{v}_3 \cdot \vec{v}_2 = \vec{v}_3 \cdot \vec{v}_1 = 0$ and $\vec{v}_3 \cdot \vec{v}_3 = 1$.

$$\hat{e}_i \hat{e}_j = \delta_{ij} \tag{2.17}$$

$$\begin{aligned}
\cos\theta &= \frac{|\vec{u}_{\parallel}|}{|\vec{u}|} \\
&= \frac{\vec{u}\cdot\hat{v}}{|\vec{u}|} \\
&= \frac{\vec{u}\cdot\vec{v}}{|\vec{u}||\vec{v}|} \\
|\vec{u}||\vec{v}|\cos\theta &= \vec{u}\cdot\vec{v}
\end{aligned} \tag{2.18}$$

M vectors can be assembled to form an $N \times M$ *matrix*

$$\mathbf{A} = [\vec{a}_1\ \vec{a}_2\ \cdots\ \vec{a}_N] = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1M} \\ a_{21} & \ddots & & & \\ \vdots & & & & \\ a_{N1} & & & & a_{NM} \end{bmatrix} \quad . \tag{2.19}$$

Matrices can likewise be assembled to form higher-dimensional *tensors* [], although we will not need them here.

The *transpose* of a matrix is

$$\mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{N1} \\ a_{12} & \ddots & & \\ \vdots & & & \\ \vdots & & & \\ a_{1M} & & & a_{MN} \end{bmatrix} \quad , \tag{2.20}$$

and the *complex conjugate* of the transpose is the *adjoint* $\mathbf{A}^\dagger = \mathbf{A}^{*T}$ where each complex element $a_{nm} = x_{nm} + iy_{nm}$ ($i^2 = -1$) is replaced by its conjugate $a^*_{mn} = x_{mn} - iy_{mn}$. If $\mathbf{A} = \mathbf{A}^\dagger$, the matrix is said to be *Hermitian*. The adjoint (or transpose) of a product is equal to the products of the adjoints (or transposes), in the opposite order:

$$\begin{aligned}
(\mathbf{AB})^T &= \sum \\
&= \mathbf{B}^T\mathbf{A}^T
\end{aligned} \tag{2.21}$$

...

If $\mathbf{A}$ is an $N \times M$ matrix and $\mathbf{B}$ is an $M \times L$ matrix, their product $\mathbf{AB}$ is an $N \times L$ matrix with elements

$$(\mathbf{AB})_{nl} = \sum_{m=1}^{M} a_{nm} b_{ml} \quad . \tag{2.22}$$

From the definition, matrix multiplication is *distributive*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \tag{2.23}$$

and *associative*

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \tag{2.24}$$

but need not be *commutative*

$$\mathbf{AB} \neq \mathbf{BA} \quad . \tag{2.25}$$

The product of a vector and a matrix

$$\mathbf{A}\vec{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1M} \\ a_{21} & \ddots & & & \\ \vdots & & & & \\ a_{N1} & & & & a_{NM} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix} \tag{2.26}$$

is a vector with elements

$$(\mathbf{A}\vec{v})_n = \sum_{m=1}^{M} a_{nm} v_m \quad . \tag{2.27}$$

The *inner product* or *dot product* of two vectors is

$$\vec{u}^T \vec{v} = [u_1 \; u_2 \; \cdots \; u_M] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix} = \sum_{m=1}^{M} u_m v_m \equiv \vec{u} \cdot \vec{v} \tag{2.28}$$

and the *outer product* is

$$\vec{u}\vec{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix} [v_1 \; v_2 \; \cdots \; v_M] = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots \\ u_2 v_1 & u_2 v_2 & \\ \vdots & & \ddots \end{bmatrix} \quad . \tag{2.29}$$

complex, norm, magnitude adjoint
positive definite matrix
basis

2.2 SYSTEMS OF EQUATIONS

If $\mathbf{A}$ is a square matrix, and $\vec{v}$ is a vector such that

$$\mathbf{A}\vec{v} = \lambda\vec{v} \quad , \tag{2.30}$$

then $\vec{v}$ is an *eigenvector* of $\mathbf{A}$ with *eigenvalue* λ. This can be rewritten as

$$(\mathbf{A} - \lambda\mathbf{I})\vec{x} = \vec{0} \quad , \tag{2.31}$$

where $\vec{0}$ is a vector of zeros. If $(\mathbf{A} - \lambda\mathbf{I})$ is invertible this gives the trivial solution

$$\vec{x} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\, \vec{0} = \vec{0} \quad . \tag{2.32}$$

Therefore, for non-trivial eigenvectors to exist $(\mathbf{A} - \lambda\mathbf{I})$ must not be invertible and hence

$$|\mathbf{A} - \lambda\mathbf{I}| \equiv \Delta(\lambda) = 0 \quad . \tag{2.33}$$

This defines the *characteristic polynomial* $\Delta(\lambda)$ for $\mathbf{A}$.

char poly imply N eigenvalues
determinant

$$V(\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_N) = 1 \tag{2.34}$$

$$V(\ldots, \vec{x}_i, \ldots, \vec{x}_i, \ldots) = 0 \tag{2.35}$$

$$\begin{aligned} V(\ldots, \vec{x}_i + \vec{x}_j, \ldots, \vec{x}_i + \vec{x}_j, \ldots) &= V(\ldots, \vec{x}_i, \ldots, \vec{x}_i, \ldots) + V(\ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots) \\ &\quad + V(\ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots) + V(\ldots, \vec{x}_j, \ldots, \vec{x}_j, \ldots) \\ &= V(\ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots) + V(\ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots) \end{aligned} \tag{2.36}$$

$$V(\ldots, \alpha\vec{x}_i + \beta\vec{x}_j, \ldots) = \alpha V(\ldots, \vec{x}_i \ldots) + \beta V(\ldots, \vec{x}_j \ldots) \tag{2.37}$$

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) \tag{2.38}$$

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})} \tag{2.39}$$

as transformation volume
recursive definition
A elements a_{ij}

$$\det \mathbf{A} \equiv |\mathbf{A}| = \sum_j a_{ij} C_{ij} \tag{2.40}$$

cofactor

$$C_{ij} = (-1)^{i+j} M_{ij} \tag{2.41}$$

minor

$$\begin{aligned} \det(\mathbf{A}) &= V\left(\sum_i a_{i1}\hat{e}_i, \sum_j a_{j2}\hat{e}_j, \ldots\right) \\ &= \sum_i a_{i1} V\left(\hat{e}_i, \sum_j a_{j2}\hat{e}_j, \ldots\right) \\ &\equiv \sum_i a_{i1} V_{i1} \\ 1 &= \frac{1}{\det(\mathbf{A})} \sum_i a_{i1} V_{i1} \end{aligned} \tag{2.42}$$

$$
\begin{aligned}
\sum_j a_{j1} V_{j2} &= \sum_j a_{j1} V \left(\sum_i a_{i1} \hat{e}_i, \hat{e}_j, \sum_k a_{k3} \hat{e}_k, \ldots \right) \\
&= V \left(\sum_i a_{i1} \hat{e}_i, \sum_j a_{j1} \hat{e}_j, \sum_k a_{k3} \hat{e}_k, \ldots \right) \\
&= 0
\end{aligned} \tag{2.43}
$$

$$
\frac{1}{\det(\mathbf{A})} \sum_i a_{ij} V_{ik} = \delta_{jk} \tag{2.44}
$$

$$
\frac{1}{\det(\mathbf{A})} \mathbf{V}^T \mathbf{A} = \mathbf{I} \tag{2.45}
$$

$$
\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \mathbf{V}^T \tag{2.46}
$$

use in change of coordinates noise problem
det prod = prod det
det = prog eigenvalues
determinant as volume

$$
\det(\mathbf{M}^{-1}) \det(\mathbf{A}) \det(\mathbf{M}) = \det(\mathbf{\Lambda}) \tag{2.47}
$$

$$
\det(\mathbf{A}) = \det(\mathbf{\Lambda}) = \prod_{n=1}^{N} \lambda_n \tag{2.48}
$$

adjoint matrix
inverse

$$
\mathbf{A}^{-1} = \frac{\mathbf{C}^T}{|\mathbf{A}|} \tag{2.49}
$$

Gauss-Jordan elimination

$$
\begin{aligned}
\mathbf{A}\mathbf{A}^{-1} &= \mathbf{I} \\
\mathbf{T}\mathbf{A}\mathbf{A}^{-1} &= \mathbf{T}\mathbf{I} \\
(\mathbf{T}\mathbf{A})\mathbf{A}^{-1} &= \mathbf{T}
\end{aligned} \tag{2.50}
$$

inverse products equals product inverses
domain and range of set of equations
If $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}$, the matrix is *unitary*.
Gauss elimination
characteristic polynomial
diagonalization

similarity transform

$$\vec{x} = \sum_{n=1}^{N} a_n \vec{v}_n = [\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_N] \begin{bmatrix} \\ \\ \\ a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \equiv \mathbf{M}\vec{a} \tag{2.51}$$

modal matrix **M**

$$\mathbf{A}\vec{x} = \vec{y} \tag{2.52}$$

range null space rank dimension range
repeated eigenvalues
Jordan blocks
cf generation of solutions to differential equations!
rank
generalized eigenvalues, eigenvectors
generalized inverse

2.3 SINGULAR VALUES

norm L^2
Hermitian adjoint
transpose
transpose product = product transpose adjoint product = product adjoint
Hermitian
symmetric

$$\vec{v}^\dagger \mathbf{H}\vec{v} = \vec{v}^\dagger \lambda \vec{v} = \lambda |\vec{v}|^2 \tag{2.53}$$

$$\left(\vec{v}^\dagger \mathbf{H}\vec{v}\right)^\dagger = \lambda^\dagger |\vec{v}|^2 \tag{2.54}$$

$$\left(\vec{v}^\dagger \mathbf{H}\vec{v}\right)^\dagger = \vec{v}^\dagger \mathbf{H}^\dagger \vec{v} = \vec{v}^\dagger \mathbf{H}\vec{v} = \lambda |\vec{v}|^2 \tag{2.55}$$

$\lambda^\dagger = \lambda$ λ real

$$\mathbf{M}^{-1}\mathbf{H}\mathbf{M} = \mathbf{\Lambda} \tag{2.56}$$

$$\begin{aligned}
\left(\mathbf{M}^{-1}\mathbf{H}\mathbf{M}\right)^{\dagger} &= \mathbf{\Lambda}^{\dagger} \\
\mathbf{M}^{\dagger}\mathbf{H}^{\dagger}\mathbf{M}^{-1\dagger} &= \mathbf{\Lambda} \\
\mathbf{M}^{\dagger}\mathbf{H}\mathbf{M}^{-1\dagger} &= \mathbf{M}^{-1}\mathbf{H}\mathbf{M}
\end{aligned} \tag{2.57}$$

$\mathbf{M}^{\dagger} = \mathbf{M}^{-1}$
orthonormal basis
positive semidefinite
positive definite
eigenvalues
$\mathbf{H} = \mathbf{A}^{\dagger}\mathbf{A}$

$$\mathbf{H}^{\dagger} = \left(\mathbf{A}^{\dagger}\mathbf{A}\right)^{\dagger} = \mathbf{A}^{\dagger}\mathbf{A}^{\dagger\dagger} = \mathbf{A}^{\dagger}\mathbf{A} = \mathbf{H} \tag{2.58}$$

$$\mathbf{H}\vec{v}_n = \lambda_n \vec{v}_n \tag{2.59}$$

rank of a product
$\sigma_i = \sqrt{\lambda_i}$

$$\vec{u}_i = \frac{1}{\sigma_i}\mathbf{A}\vec{v}_i \tag{2.60}$$

$$\begin{aligned}
\vec{u}_i^{\dagger}\vec{u}_j &= \frac{1}{\sigma_i\sigma_j}\vec{v}_i^{\dagger}\mathbf{A}^{\dagger}\mathbf{A}\vec{v}_j \\
&= \frac{1}{\sigma_i\sigma_j}\vec{v}_i^{\dagger}\mathbf{H}\vec{v}_j \\
&= \frac{\lambda_j}{\sigma_i\sigma_j}\vec{v}_i^{\dagger}\vec{v}_j \\
&= \delta_{ij}
\end{aligned} \tag{2.61}$$

$$\vec{u}_i^{\dagger}\mathbf{A}\vec{v}_j = \vec{u}_i^{\dagger}\sigma_j\vec{u}_j = \sigma_j\delta_{ij} \tag{2.62}$$

extend by Gram-Schmidt orthogonalization
v beyond rank in null space

$$\mathbf{U}^{\dagger}\mathbf{A}\mathbf{V} = \mathbf{\Sigma} \tag{2.63}$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & & & & 0 \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{bmatrix} \tag{2.64}$$

$$\mathbf{A} = \mathbf{U\Sigma V}^\dagger \tag{2.65}$$

SVD Singular Value Decomposition
rank
SVD for full row rank simplification relate to Zen Kalman filter

2.4 FUNCTIONS OF MATRICES

matrix function power series converges for eigenvalues

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2t^2 + \cdots \tag{2.66}$$

$$\begin{aligned} \frac{d}{dt}e^{\mathbf{A}t} &= \mathbf{A} + \mathbf{A}^2t + \cdots \\ &= \mathbf{A}e^{\mathbf{A}t} \end{aligned} \tag{2.67}$$

$$\begin{aligned} e^{\mathbf{A}t}e^{\mathbf{B}t} &= \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2t^2 + \cdots\right)\left(\mathbf{I} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2t^2 + \cdots\right) \\ &= \mathbf{I} + (\mathbf{A} + \mathbf{B})t + \frac{1}{2}\mathbf{A}^2t^2 + \frac{1}{2}\mathbf{B}^2t^2 + \mathbf{AB}t + \cdots \end{aligned} \tag{2.68}$$

$$\begin{aligned} e^{(\mathbf{A}+\mathbf{B})t} &= \mathbf{I} + (\mathbf{A} + \mathbf{B})t + \frac{1}{2}(\mathbf{A} + \mathbf{B})^2t^2 + \cdots \\ &= \mathbf{I} + (\mathbf{A} + \mathbf{B})t + \frac{1}{2}\mathbf{A}^2t^2 + \frac{1}{2}\mathbf{B}^2t^2 + \frac{1}{2}(\mathbf{AB} + \mathbf{BA})t + \cdots \end{aligned} \tag{2.69}$$

Cayley-Hamilton
relates $\mathbf{M}^n$ to lower order powers
mult char poly $\mathbf{M}^{-1}$ to get inverse from eigenvalues

$$\begin{aligned} 0 &= |\lambda\mathbf{I} - \mathbf{A}| \\ &= \lambda^N + a_{N-1}\lambda^{N-1} + \cdots + a_0 \\ &\equiv C(\lambda) \end{aligned} \tag{2.70}$$

$$a_0 = |\mathbf{A}| \tag{2.71}$$

$$\begin{aligned}\mathbf{0} &= C(\mathbf{A}) \\ &= \mathbf{A}^N + a_{N-1}\mathbf{A}^{N-1} + \cdots + a_0\mathbf{I}\end{aligned} \tag{2.72}$$

$$\mathbf{A} = \mathbf{M\Lambda M} \tag{2.73}$$

$$\begin{aligned}\mathbf{A}^2 &= \mathbf{M\Lambda M}^{-1}\mathbf{M\Lambda M}^{-1} \\ &= \mathbf{M\Lambda\Lambda M}^{-1} \\ &= \mathbf{M\Lambda}^2\mathbf{M}^{-1}\end{aligned} \tag{2.74}$$

$$\begin{aligned}C(\mathbf{A}) &= \mathbf{A}^N + a_{N-1}\mathbf{A}^{N-1} + \cdots + a_0\mathbf{I} \\ &= \mathbf{M}\left(\mathbf{\Lambda}^N + a_{N-1}\mathbf{\Lambda}^{N-1} + \cdots + a_0\mathbf{I}\right)\mathbf{M}^{-1} \\ &= \mathbf{M0M}^{-1} \\ &= \mathbf{0}\end{aligned} \tag{2.75}$$

generalize to Jordan blocks
inverse

$$\begin{aligned}\mathbf{0} &= \mathbf{A}^N + a_{N-1}\mathbf{A}^{N-1} + \cdots + a_0\mathbf{I} \\ &= \mathbf{A}^{N-1} + a_{N-1}\mathbf{A}^{N-2} + \cdots + a_0\mathbf{A}^{-1} \\ \mathbf{A}^{-1} &= -\frac{1}{a_0}\left(\mathbf{A}^{N-1} + a_{N-1}\mathbf{A}^{N-2} + \cdots + a_1\mathbf{I}\right)\end{aligned} \tag{2.76}$$

exponential

$$\begin{aligned}e^{\mathbf{A}t} &= \mathbf{I} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2t^2 + \cdots \\ &= \alpha_0(t)\mathbf{I} + \alpha_1(t)\mathbf{A} + \alpha_2(t)\mathbf{A}^2 + \cdots + \alpha_{N-1}(t)\mathbf{A}^{N-1}\end{aligned} \tag{2.77}$$

functions of a matrix
using generalized inverses
go through book to move linear algebra definitions to here
drop use of dots for matrix multiplication?
ODE chapter problem on exponential matrix solution

2.5 SELECTED REFERENCES

[Golub & Loan, 1989] Golub, G.H., & Loan, C.F. Van. (1989). *Matrix Computations*. 2nd edn. Baltimore, MD: Johns Hopkins University Press.

2.6 PROBLEMS

(2.1) pseudo-inverse from SVD
(2.2) rotation unitary, dot product angle
(2.3) exponential angular momentum = rotation (lead into SU(2) problem)