Linear Algebra

Hugo Rivera

Math 254, NMT. Summer 2015

Contents

List of Theorems

1 Linear Equations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(x) = Ax$, A is an $m \times n$ matrix, and $Ax = b$ is a linear system. Logically equivalent:

- T is one-to-one (at most one solution for all vectors b).
- None of the columns of A is a linear combination of the others.
- $Ax = 0$ has only the trivial solution.
- The columns of A are linearly independent.

Logically equivalent:

- T is onto (at least one solution for all vectors b).
- The columns of A span $\mathbb{R}^m$.
- For all b, b is a linear combination of the columns of A.
- For all b, $Ax = b$ has a solution.
- A has a pivot in every row.

1.1 Systems of Linear Equations

linear equation equation that can be written as $a_1x_1 + \cdots + a_nx_n = b$. Numbers a_i are **coefficients**.

linear system many linear equations.

solution A number making each equation a true statement on substitution of the x_i variables.

solution set all possible solutions of a linear system

equivalence of linear systems: same solution set? equal linear systems.

consistent has one or infinitely many solutions. **Inconsistent** has none.

augmented matrix matrix representation for linear system. Rows are equations, columns variables.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right]$$

$m \times n$ **matrix** has m rows, n columns

solving linear system replace system with an equivalent, but easier to solve.

elementary row operations all reversible.

1. replace one row by the sum of itself and a multiple of another row
2. interchange two rows
3. multiply row by non-zero constant

row equivalence if two matrices can be transformed into each other through row ops.

existence and uniqueness questions fundamental question. Does a system have a solution, is it the only one?

1.2 Row Reduction and Echelon Forms

leading entry leftmost nonzero entry of a nonzero row

echelon form (rectangular matrix)

1. nonzero rows above all-zero rows
2. leading entry is to the right of leading entry above
3. entries in a column below leading entry are zero

reduced row echelon form AKA reduced echelon form. Must be in echelon form and:

1. leading entry in each nonzero row is 1
2. leading 1 is the only nonzero in column

row reduction is possible for any nonzero matrix

pivot positions locations of the leading 1s in the RREF of a matrix. **Pivot columns** are the columns with such 1s.

observations one for each row

variables one for each column, excluding the right-most of the augmented matrix

basic variables correspond to pivot columns

free variables are not pivot columns. May take any value.

general solution explicit description of all solutions, such as

$$S = \begin{cases} x_1 & = 1 + 5x_3 \\ x_2 & = 4 \\ x_3 & \text{is free} \end{cases}$$

row reduction solving strategy

1. Write augmented matrix
2. Row reduction algorithm to obtain echelon form
3. If consistent, continue. Otherwise no solution.
4. Continue row reduction until RREF
5. Write system of equations
6. Rewrite nonzero equations so basic variables are expressed in terms of free variables.

Th. 1.1 (Uniqueness of RREF) *Each matrix is row equivalent to one and only one reduced echelon matrix.*

Th. 1.2 (Existence and Uniqueness) *A linear system is consistent iff the rightmost column of the augmented matrix is not a pivot column. Iff the aug. matrix has no row of form $[0, \dots 0, b]; b \neq 0$.*

If the linear system is consistent, then it has a unique solution when there are no free variables and infinitely many solutions if there are.

1.2.1 Row reduction algorithm

1. Begin with the leftmost nonzero column, now a pivot column. Pivot position is at the top.
2. Select nonzero as a pivot. Interchange rows as needed.
3. Row replacement operations to create zeros below the pivot.
4. Ignore pivot position's row and all above, pick a new row and repeat 1-3 until no more nonzero rows to modify.
5. Make all pivots 1 by scaling operations.

1.3 Vector Equations

list of numbers is an intuitive definition of a vector (until chapter 4). Example: $(1, 2, 3)$.

column vector ordered arrays represented by $n \times 1$ matrix.

scalar multiplication multiply each entry in the vector by the scalar

vector sum add corresponding entries

zero vector all entries are zero

geometric interpretation a point in n-dimensional space. Can also be an arrow from the origin to that point.

parallelogram addition rule u + v corresponds to the fourth vertex of the parallelogram with vertices **u, v, 0**

Linear combination The vector $\mathbf{y}$ is a linear combo of the vectors $\mathbf{v_i}$ given the scalars (or **weights**) c_i if:

$$\mathbf{y} = c_1\mathbf{v_1} + \dots + c_p\mathbf{v_p}$$

$\mathbf{y}$ **can be generated** by a linear combination $V = \mathbf{y}$ only if the augmented matrix $[V\ \mathbf{y}]$ has a solution.

Span of vectors Span$\{\mathbf{v_1} \dots \mathbf{v_p}\}$ is called the subset of $\mathbb{R}^n$ spanned or generated by the vectors. All vectors that can be written with scalars $c_1 \dots c_p$ as:

$$c_1 v_1 + \dots + c_p v_p$$

Is a vector b in a span is tantamount to asking: does the vector equation $x_1 v_1 + \dots + x_p v_p = b$ have a solution?

Always in a span the multiples of v_i, and the zero vector

Span{v} is a **line** with all scalar multiples of **v** or the **origin** if $v = 0$.

Span{u, v} is a **plane** containing **u, v, 0** when u and v are not scalar multiples of each other and when $u \neq 0$ and $v \neq 0$.

Algebraic properties of vectors in $\mathbb{R}^n$ space, let c, d be scalars and u, v be vectors in:

1. u + v = v + u
2. (u + v) + w = u + (v + w)
3. u + 0 = 0 + u = u
4. u + (-u) = -u + u = 0
5. -u = (-1)u
6. c(u + v) = cu + cv
7. (c + d)u = cu + du
8. c(du) = (cd)u
9. 1u = u

1.4 Matrix Equation Ax = b

fundamental idea linear combination of vectors is the product of a matrix and a vector. Can rephrase concepts of section 1.3.

matrix × vector linear combination of the columns using the corresponding entries of the vector as weights. Only defined if columns of A = entries in **x**.

$$A\mathbf{x} = x_1\mathbf{a_1} + \dots x_n\mathbf{a_n}$$

matrix equation has form A**x = b** where A is a matrix, x,b are matrices.

Existence of solutions $A\mathbf{x} = \mathbf{b}$ has a solution iff **b** is a linear combination of the columns of A.

Row-vector rule for matrix-vector product A**x**. If it is defined, then the ith entry in the result is the sum of the products of corresponding entries from row i of A and from vector **x**.

$$A\mathbf{x} = \begin{bmatrix} \sum a_{1i}x_1 \\ \sum a_{2i}x_2 \\ \dots \\ \sum a_{ni}x_n \end{bmatrix}$$

Identity matrix has 1s on diagonal and 0s everywhere else. I**x = x** for every $\mathbf{x} \in \mathbb{R}^n$

Proof of theorem 1.4 statements $1, 2, 3$ have been shown to be true; now assume statement 4 is false. Then the augmented matrix U would be inconsistent and A**x = b** would have no solution. If statement 4 is true, then the system is consistent and has at least one solution.

Th. 1.3 (Notation) *If A is an $m \times n$ matrix and $b \in \mathbb{R}^n$ then the matrix equation, vector equation and augmented matrix share the same solution set:*

$$A\mathbf{x} = \mathbf{b}$$

$$x_1\mathbf{a_1} + \cdots + x_n\mathbf{a_n} = \mathbf{b}$$

$$\left[\begin{array}{ccc|c} \mathbf{a_1} & \cdots & \mathbf{a_n} & \mathbf{b} \end{array}\right]$$

Th. 1.4 (Span, linear combo., pivots) *Let A be an $m \times n$ matrix. The following statements are logically equivalent:*

1. *For each $\mathbf{b} \in \mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*
2. *Each $\mathbf{b} \in \mathbb{R}^n$ is a linear combination of the columns in A.*
3. *The columns of A span $\mathbb{R}^n$.*
4. *A has a pivot position in every row. That is the RREF of A (not the augmented matrix) has a pivot in every row.*

Th. 1.5 (Algebraic properties of matrix $\times$ vector) *Let A be an $m \times n$ matrix, u, v be n length vectors, and c is a scalar.*

1. *$A(u + v) = Au + Av$*
2. *$A(cu) = c(Au)$*
3. *uA is undefined*

1.5 Solution Sets of Linear Systems

homogeneous systems have $\mathbf{b} = \mathbf{0}$.

trivial solution is always 0.

nontrivial solution the system has a free variable.

Span of vectors $\{v_i\}$ can be used to represent a solution set of $Ax = 0$ using the right v_i vectors. The set Span$\{0\}$ represents systems with only the trivial solution.

1 free variable or more iff there's a nontrivial solution.

parametric vector equation solution set described using free variables as parameters. Solution is a linear combination of the vectors.

$$\mathbf{x} = s\mathbf{u} + t\mathbf{v} \quad (s, t \in \mathbb{R})$$

parametric vector form parametric vector equation with the vectors $\mathbf{u} \cdots$ written explicitly.

solution of non-homogeneous Any of the homo. solutions: v_h, then nonhomo. solutions: $v_h + p$ where p is any solution of the system. A non-homogeneous system has the same solution set, just translated by vector p. This only applies if nonhomo. system has at least one solution.

vector translation adding a vector, think of moving it around.

writing a solution set of a consistent system in parametric vector form

Th. 1.6 (solution set of non/homogeneous) *Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent and let $\mathbf{p}$ be a solution. Solution set of is the set of all vectors of form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$, where $\mathbf{v_h}$ is any solution of $A\mathbf{x} = \mathbf{0}$. This only applies if the system has nontrivial solutions.*

1.6 Applications

homogeneous system in economics equilibrium between input and output. Quick way to turn input-output table into matrix: identity matrix - IO table.

Leontief exchange model simpler version of the production model. No demand vector. Assumed that everything produced is consumed by the "productive" sectors of the economy.

Example

balancing chemical equations yup

network flow write equations: sum of flow in = sum of flow out

1.7 Linear Independence

independence a set of p vectors in $\mathbb{R}^n$ space is linear independence if $x_1\mathbf{v_1} \ldots x_p\mathbf{v_p} = \mathbf{0}$ or $Ax = 0$ has only the trivial solution.

dependent if non-zero weights c_i such that $c_1\mathbf{v_1} \ldots c_p\mathbf{v_p} = \mathbf{0}$.

linear combination a linear dependence represents a nontrivial solution: $\mathbf{v_p} = (c_1\mathbf{v_1} \ldots c_{p-1}\mathbf{v_{p-1}})/c_p$

linear dependence relation called so only when all weights c_i are non-zero.

linear independence of columns in a matrix happens only when $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

one vector is linear independent iff it is not the zero vector.

two vectors are linearly dependent when they are multiples of each other.

proof of theorem 7 Assume a vector is a linear combination of the others. Subtract it to produce a homogeneous equation with at least one nonzero coefficient -1. Thus the vectors are linearly dependent. Now assume the set is linearly independent. Then the homogeneous equation can be rewritten by subtracting one of the nonzero vectors and dividing its coefficient to get a linear combination. QED

proof of theorem 8 if there are more columns than rows, then there must be a free variable. This means $Ax = 0$ has a nontrivial solution.

proof of theorem 9 consider the equation with $v_1 = 0$, $1v_1 + 0v_2 \cdots + 0v_p = 0$. It must be linearly dependent.

Th. 1.7 (Characterization of Linearly Dep. Sets) *A linearly dependent set has* **at least one** *vector that is a linear combination of the others.*

Th. 1.8 (More vectors than vector entries) *Must be dependent.*

Th. 1.9 (Contains Zero vector) *Must be dependent.*

1.8 Linear Transformations

transformation, function, mapping a rule that assigns one vector from $\mathbb{R}^n$ to another in $\mathbb{R}^m$, where n may equal m.

domain set of all inputs

codomain subset of outputs

range all possible outputs, a.k.a. **image**.

matrix transformation transforms a vector from $\mathbb{R}^n$ to one from $\mathbb{R}^m$ through matrix multiplication $x \mapsto Ax$.

linear transformation a transformation T is linear (and can be represented by matrix multiplication) iff $T(u+v) = T(u) + T(v)$ and $T(cu) = cT(u)$ for vectors u, v and the scalar c.

$T(0) = 0$ **follows** because $T(0) = T(0u) = 0T(u) = 0$.

1.9 Matrix of a Linear Transformations

standard matrix for a linear trans. all linear transformations can be defined using a standard matrix A like this: $T(x) = Ax$. If $T : \mathbb{R}^n \to \mathbb{R}^m$ then A is $m \times n$.

solving for standard matrix e_i is the ith column of the identity matrix. $A = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}$

geometric linear trans. think of performing the $\mathbb{R}^2$ transformation on a unit square because T is solely defined by its actions on the identity matrix I.

onto mappings T is maps onto $\mathbb{R}^m$ if every b in $\mathbb{R}^m$ is the image of **at least one** $x \in \mathbb{R}^n$

one-to-one mappings T is one-to-one if every b in $\mathbb{R}^m$ is the image of **at most one** $x \in \mathbb{R}^n$

proof of theorem 11 Assume T is one-to-one. Then $T(0) = 0$ counts as one solution and only the trivial solution. Now assume T is not one-to-one. Then there exist at least two distinct vectors $T(u) = b = T(v)$. Because they are distinct, subtracting them gives another solution to $T(x) = 0$, $T(u - v) = b - b = 0$.

proof of theorem 12 (a) The columns of A span $\mathbb{R}^m$ iff $Ax = b$ is consistent for every b. In other words, if $T(x) = b$ has at least one solution. Thus T is onto.

(b) So T is one-to-one iff $Ax = 0$ has only the trivial solution. This happens iff the columns of A are linearly independent as was noted previously.

Th. 1.10 (Linear trans., unique matrix) *If $T : \mathbb{R}^n \to \mathbb{R}^m$ then there exists a unique $m \times n$ matrix A such that $T(x) = Ax$. This matrix is defined as $A = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}$*

Th. 1.11 (One-to-one, trivial) *T is one-to-one iff $T(x) = 0$ has only the trivial solution.*

Th. 1.12 (Linear trans. and standard matrix) *$T : \mathbb{R}^n \to \mathbb{R}^m$ and A is an $m \times n$ matrix. (a) T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ iff the columns of A span $\mathbb{R}^m$.*

(b) T is one-to-one iff the columns of A are linearly independent.

Some super fun linear transformations

rotation

reflection across x_1 axis would be $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

contraction/expansion scale by a factor of k: $\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$

shears stretches the "top" more than the "bottom," or vice-verse (slant-like) $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

projections a lot like $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

1.10 Examples of Linear Models

Kirchhoff's law see below

DC networks sum of all voltages in a loop = sum of all (resistors * adjacent current). Flows from positive to negative, otherwise negate the voltage.

difference equations $x_{k+1} = Ax_k$ a recurrence relation.

2 Matrix Algebra

2.1 Matrix Operations

matrix has m rows and n columns, denoted as $m \times n$. The element in the ith row and jth column is represented by A_{ij}.

zero matrix all entries are zero

identity matrix all entries on the diagonal are one and the rest are zero.

sane addition and scalar multiplication rules for two matrices A, B and two scalars r, s.

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

proofs by decomposing matrices into scalars A_{ij} or column vectors a_i. Show entries are equal and matrices are same size.

multiplication goal could have element-wise multiplication, but it is better to have $A(Bx) = (AB)x$, composable matrix-vector multiplication.

matrix multiplication only works for two matrices, A of size $m \times n$ and B of size $n \times p$. The result has size $n \times p$. Defined as $AB = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$.

row-column rule for computing AB. The (i, j) entry in AB is $(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$ where n is the shared size.

WARNINGS

Not commutative $AB \neq BA$.

No cancellation laws $\neg(AB = AC \implies B = C)$ unless A is invertible.

Zero is not the only matrix that can create zero $AB = 0 \not\Longrightarrow A = 0 \vee B = 0$.

powers $A^k = A \cdots A$, the matrix A multiplied by itself k times. $A^0 = I$ so that $A^0 x = x$.

transpose A^T is the $n \times m$ transpose of the $m \times n$ sized matrix A. The entries are defined: $A^T_{ij} = A_{ji}$.

Th. 2.1 (Matrix Multiplication rules) *Let A be an $m \times n$ matrix and let B, C have sizes for which these sums and products are defined.*

1. *associative:* $(AB)C = A(BC)$
2. *left distributive:* $A(B + C) = AB + AC$
3. *right distributive:* $(B + C)A = BA + CA$
4. $r(AB) = (rA)B = A(rB)$ *for any scalar* r
5. *identity elements:* $I_m A = A = AI_n$.

Th. 2.2 (Transpose rules) *Let A, B be matrices whose sizes are appropriate for the following sums and products.*

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. *For any scalar* $r, (rA)^T = rA^t$
4. $(AB)^T = B^T A^T$

2.2 Inverse of a Matrix

inverse of a matrix A^{-1} is a unique element which exists only for certain square matrices. Properties: $AA^{-1} = A^{-1}A = I$.

invertible matrices have an inverse. Their determinant is not zero. Also called nonsingular.

noninvertible matrices are called singular or degenerate.

inverse of $\mathbb{R}^2$ matrix is easy to calculate

proof of Th. 2.4 Take any $b \in \mathbb{R}^n$. The solution $x = A^{-1}b$ exists. Simply substitute to see. x is the only solution. Let u be another arbitrary vector. If $Au = b$, then $Iu = A^{-1}b$ and $u = A^{-1}b$.

Row reduction is faster than finding the inverse of a matrix. Except, perhaps, in 2×2 matrices.

Elementary matrix E, an identity matrix after a single elementary row operation.

Inverse of elementary matrix is the same matrix that transforms E back into I. Generalized into Th. 2.6

Algorithm for finding the inverse of A. Row reduce the matrix $\begin{bmatrix} A & I \end{bmatrix}$. If the result has form $\begin{bmatrix} I & A^{-1} \end{bmatrix}$, then A has an inverse. Otherwise, there is none.

Alternative algorithm for finding one or two columns of A^{-1}. The columns of A^{-1} are the solutions of the system $Ax = e_1; Ax = e_2; \cdots Ax = e_n$. So, to find the pth column of the inverse, simply solve $Ax = e_p$.

Th. 2.3 (Inverse of 2 by 2) *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If its determinant* $ad - bc \neq 0$, *then the inverse of A is* $A^{-1} = (1/(ad-bc)) \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. *If the determinant is zero, then A is not invertible. The* **determinant** *of this matrix is* $\det A = ad - bc$.

Th. 2.4 (Solutions if invertible) *If A is an invertible $n \times n$ matrix, then for each $b \in \mathbb{R}^n$, the equation $Ax = b$ has the unique solution $x = A^{-1}b$.*

Th. 2.5 (Rules and the Transpose) *Assume A is invertible. Assume B is invertible and the same size as A.*

1. A^{-1} *is invertible.*
2. $(A^{-1})^{-1} = A$.
3. $(AB)^{-1} = B^{-1}A^{-1}$.
4. $(A^T)^{-1} = (A^{-1})^T = A^{-T}$.
5. *The product of $n \times n$ invertible matrices is invertible and it is the product of their inverses in the reverse order.*

Th. 2.6 (Inverse of a matrix) *An $n \times n$ matrix A is invertible iff A is row equivalent to I_n. The same sequence that transforms A to I_n also transforms I_n to A^{-1}.*

2.3 Characterizations of Invertible Matrices

Th. 2.7 (Invertible Matrix Theorem) *A is a square $n \times n$ matrix. These statements are logically equivalent:*

1. *A is an invertible matrix.*
2. *A is row equivalent to I_n.*
3. *A has n pivot columns.*
4. *The equation $Ax = 0$ has only the trivial solution.*
5. *The columns of A form a linearly independent set.*
6. *The linear transformation $x \mapsto Ax$ is one-to-one.*
7. *The equation $Ax = b$ has at least one solution for each $b \in \mathbb{R}^n$.*
8. *The columns of A span $\mathbb{R}^n$.*
9. *The linear transformation $x \mapsto Ax$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.*
10. *There is a pair of matrices C, D such that $CA = I_n$ and $AD = I_n$.*
11. *A^T is an invertible matrix.*

IMT separates square matrices into two disjoint classes.
Only square matrices can be treated with the IMT
Fact if $AB = I_n$ then A, B are both invertible with $B = A^{-1}; A = B^{-1}$.
Singular, degenrate matrices are not invertible.

Th. 2.8 (Invertible Transformation) *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ with standard matrix A. Then T is invertible iff A is an invertible matrix. If so, then $S(x) = A^{-1}x$ satisfies $T(S(x)) = x$ and $S(T(x)) = x$.*

2.4 Partitioned Matrices

Matrices can be split into chunks to alleviate the burden of multiplying large matrices.

2.5 Matrix Factorization

Matrices can be factored into a pair of upper and lower triangular matrices $A = LU$.

2.6 Leontief Input-Output Model

Production vector x lists the outputs of each of n sectors of an economy
Unit consumption vector for each sector, lists the inputs from other sectors to produce one unit of output.
Consumption matrix C, is a matrix composed of all the unit cons. vectors.
Final demand d, lists the amount of goods demanded by the non-productive part of economy.
Production equation $x = Cx + d$ or $(I - C)x = d$ or if $I - C$ is invertible, $x = (I - C)^{-1}d$.
Modification If the demand changes to $d + \Delta d$, then the economy must shift by $\Delta x = (I - C)^{-1}\Delta d$.
Formula for $(I - C)^{-1}$ is similar to the geometric series: $(I - C)^{-1} = \sum_{n=0}^{\infty} C^n$.
Approximation If $C^m \to 0$ quickly enough, then this approximation applies: $(I - C)^{-1} \approx \sum_{n=0}^{m} C^n$.
Example manufacting is c_1, big agri is c_2, and services is c_3. If it takes 50 units from other parts of manufacturing, 20 units from agriculture, and 10 units from services to produce 100 units of manufacturing then:

$$c_1 = \begin{bmatrix} 0.5 \\ 0.2 \\ 0.1 \end{bmatrix} \quad C = \begin{bmatrix} 0.5 & 0.4 & 0.2 \\ 0.2 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{bmatrix} \quad d = \begin{bmatrix} 50 \\ 30 \\ 20 \end{bmatrix}$$

This system can be solved to get $x = \begin{bmatrix} 226 \\ 119 \\ 78 \end{bmatrix}$.

2.7 Applications to Computer Graphics

So cool! 3D to 2D projection * rotation * translation * scale * crappy toyota vertices $\mapsto$ pretty toyota vertices

3 Determinants

3.1 Introduction to Determinants

Simple Determinants a 2×2 matrix has determinant $ad - bc = a_{11}d_{22} - b_{12}c_{21}$. A 1×1 matrix has determinant a_{11}
Deletion of a row i and a column j of a matrix is represented as A_{ij}
Cofactor along the ith row and jth columnt is $C_{ij} = (-1)^{i+j} \det A_{ij}$

Th. 3.1 (Cofactor Expansion) *Let A be an $n \times n$ matrix. The determinant can be found in two similar ways:*
Down column j is, we have $\det A = \sum_{k=1}^{n} a_{kj}C_{kj}$.
Along row i is, we have $\det A = \sum_{k=1}^{n} a_{ik}C_{ik}$.

Th. 3.2 (Triangular matrix) *A triangular matrix has a determinant equal to the product of the diagonal elements: $\det A = \prod a_{ii}$.*

3.2 Properties of Determinants

Th. 3.3 (Row Operations and Det) *Let A be a square matrix.*

1. *If a multiple of one row of A is added to another row to produce matrix A', then $\det A = \det A'$.*
2. *If two rows are interchanged to produce matrix A', then $-\det A = \det A'$.*
3. *If one row is multiplied by k to produce a matrix A', then $k \det A = \det A'$.*

Common use of Th 3.3 factor out a multiple of a row to simplify finding the determinant.

Row op based formula for calculating determinant of a matrix A row equivalent to a matrix U in echelon form. Let r be the number of row exchanges it takes to transform A. Efficient!

$$\det A = \begin{cases} (-1)^r \cdot \prod \text{ pivots in U} & \text{if A is invertible} \\ 0 & \text{if A is not invertible} \end{cases}$$

Th. 3.4 (IMT and non-zero Det) *A square matrix A is invertible iff $\det A \neq 0$.*

Th. 3.5 (Column operations and Det) *For a square matrix A, $\det A^T = \det A$.*

Th. 3.6 (Multiplication and Det) *If A and B are square matrices, then $\det AB = \det A \det B$.*

WARNING: no analogue for summation of matrices $\det(A + B)$.

Determinant is a Linear function for a certain set of matrices. Useful in advanced courses. Assume a matrix has constant columns except for one. This varying column is the parameter x to a function $T(x) = \det \begin{pmatrix} a_1 \cdots x \cdots a_n \end{pmatrix}$. It can be shown that $T(cx) = cT(x)$ and $T(u + v) = T(u) + T(v)$.

3.3 Cramer's Rule, Volume, Linear Transformations

Th. 3.7 (Cramer's Rule) *Let A be an invertible $n \times n$ matrix. For any $b \in \mathbb{R}^n$ the unique solution x of $Ax = b$ has entries given by*

$$x_i = \det A_i(b) \det A, \qquad i = 1, \ldots, n$$

Proof of Th 3.7 If $Ax = b$, then $A \cdot I_i(x) = \begin{bmatrix} Ae_1 & \cdots & Ax & \cdots & Ae_n \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & b & \cdots & a_n \end{bmatrix} = A_i(b)$. By the multiplicative property of determinants, $\det A \det I_i(x) = \det A_i(b)$. But $\det I_i(x) = x$ because it is a diagonal matrix. Because A is invertible, $\det A \neq 0$ and it follows that $x = \det A_i(b) / \det A$.

Adjugate or classical adjoint of an $n \times n$ matrix A is the matrix formed by taking the cofactors of every element in A and then transposing the resulting matrix:

$$adj A = \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}^T$$

Th. 3.8 (Inverse formula) *Let A be an invertible $n \times n$ matrix. Then*

$$A^{-1} = \frac{1}{\det A} adj A$$

Th. 3.9 (Simple area and volume) *If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. The volume of a parallelepiped (3D) determined by the columns of A is also $|\det A|$.*

Th. 3.10 (Linear trans. and area or vol.) *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation determined by a matrix A. If S is a parallelogram in $\mathbb{R}^2$, then*

$$\textit{area of } T(S) = |\det A| \cdot \textit{area of } S$$

Likewise for a parallelepiped S with everything in $\mathbb{R}^3$:

$$\textit{volume of } T(S) = |\det A| \cdot \textit{volume of } S$$

Example area of parallelogram with vertices at points $(-2, -2), (0, 3), (4, -1), (6, 4)$. First translate it to the origin $(0, 0), (2, 5), (6, 1), (8, 6)$. The area of this figure is

$$\det \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix} = 28$$

Proof of Th 3.9

Proof of Th 3.10

Generalization of Th 3.10 This theorem holds for any region with finite area in $\mathbb{R}^3$ or finite volume in $\mathbb{R}^3$.

Example with ellipse which has equation $x^2/a^2 + y^2/b^2 \leq 1$. It can be transformed into a unit sphere by the transformation with standard matrix $\left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$ with determinant ab. Because a unit sphere has volume π, the ellipse has volume $ab\pi$.

Circle Equation of a circle $(x - a)^2 + (y - b)^2 = \sqrt{r}$ which crosses three points $(x_1, y_1), \ldots, (x_3, y_3)$ is given by:

$$\det \begin{bmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{bmatrix} = 0$$

Geometric equations can be expressed in terms of determinants. Similar expressions work for general equations of lines of any dimension, spheres, cones, etc. ...

4 Vector Spaces

4.1 Vector Spaces and Subspaces

Th. 4.1 (Vector Space Axioms) *A nonempty set V with two binary operations called addition and multiplication by scalars (in this case, real numbers). Axioms must hold for all $u, v, w \in V$ and $c, d \in \mathbb{R}$:*

1. $u + v \in V$
2. $u + v = v + u$
3. $u + (v + w) = (u + v) + w$
4. *There exists a zero vector such that* $u + 0 = u$
5. *There exist inverse vectors* $-v$ *such that* $v + (-v) = 0$
6. $cu \in V$
7. $c(u + v) = cu + cv$
8. $(c + d)u = cu + du$
9. $c(du) = (cd)u$
10. *There exists a vector such that* $1u = u$

Many properties can be shown using the ten axioms

Negative shorthand $-u = (-1)u$.

Examples polynomials of n-degree $\mathbb{P}_n$, $\mathbb{R}^n$, $\mathbb{Z}_n$, continuous functions

Subspace of a vector space V is a subset H of V with:

1. $0 \in H$
2. For all $u, v \in H$, $u + v \in H$
3. For every scalar c and vector $v \in H$, $cv \in H$

Subspaces are also vector spaces

Zero subspace $\{0\}$ a trivial subspace of any vector space

Example subspaces polynomials of degree $\leq n$ in vector space $\mathbb{P}_n$, polynomials are a subspace of the vector space formed by continuous functions

Linear combination of vectors in vector space V is $c_1v_1 + c_2v_2 \in V$ for any scalars c_1, c_2.

Span of a set of n vectors is the set formed by taking all scalars $c_1, \ldots, c_n$ and creating linear combinations of vectors Span $\{v_1, \ldots, v_n\} = c_1v_1 + \cdots + c_nv_n$.

Synonyms span of vectors is the **subspace spanned or generated** by the vectors. A **generating set** for a subspace is a set of vectors such that they span the subspace.

Th. 4.2 *If $v_1, \ldots, v_n$ are in a vector space V then Span $\{v_1, \ldots, v_n\}$ forms a subspace of V.*

4.2 Null Spaces, Column Spaces, Linear Transformations

Subspace interpretation either arise as solutions to a system of homogeneous linear equation, or as the set of linear combinations of some vectors.

Null space of an $m \times n$ matrix A is the set of all solutions to $Ax = 0$. Dynamic description: all $x \in \mathbb{R}^n$ mapped into $0 \in \mathbb{R}^m$ by the mapping $x \mapsto Ax$. **Implicitly** defined. Nul $A = \{x | x \in \mathbb{R}^n$ and $Ax = 0\}$.

Th. 4.3 (Null space is subspace) *The null space of an $m \times n$ matrix is a subspace of $\mathbb{R}^n$.*

Proof of Th. 4.3

Explicit description of a null space is found by row reducing A and finding the general solution to $Ax = 0$.

Solution set formed by a spanning set is automatically linearly inependent because free variables are the weights on the vectors.

Number of free variables is the number of vectors in the null space.

Column space of an $m \times n$ matrix A is the set of all linear combinations of the columns of A. Col $A = Span\{a_1, \ldots, a_n\}$.

About null space :

1. A subspace of $\mathbb{R}^n$
2. Implicitly defined
3. Easy to tell if vector is present, check $Ab = 0$
4. Equal to $\{0\}$ iff $Ax = 0$ has only the trivial solution.

About column space :

1. A subspace of $\mathbb{R}^m$
2. Explictly defined
3. Hard to tell if vector is present, must row reduce $[A \quad v]$
4. Equal to $\mathbb{R}^m$ iff $Ax = b$ has a solution for every $b \in \mathbb{R}^m$.

Linear transformation T from a vector space V to a vector space W assigns each vector $x \in V$ to a vector $T(x) \in W$. Linear if: $T(u + v) = T(u) + T(v)$ and if $T(cu) = cT(u)$ for all vectors $u, v \in V$ and all scalars $c \in \mathbb{R}$.

Kernel of a transformation T is set of all vectors in V that map to $0 \in W$. Solutions to $T(x) = 0$.

Range of a transformation T is set of all vectors in W with form $T(x)$ for some $x \in V$.

Example differentiation is a linear transformation of real-valued continuous functions.

4.3 Linearly Independent Sets; Bases

Th. 4.4 *A set of two or more vectors with $v_1 \neq 0$ is linearly dependent iff some v_j with $j > 1$ is a linear combination of the other vectors.*

Linear independence Vectors v are linearly independent if only $c_i \neq 0$ can satisfy the equation: $c_1 v_1 + \cdots + c_p v_p = 0$.

Linear dependence relation formed when $c_i \neq 0$ for some c_i.

Basis for a subspace H of a vector space B is formed by the set $\mathcal{B} = \{b_1, \ldots, b_p\}$ if

1. $\mathcal{B}$ is a linearly independent set
2. Span $\mathcal{B} = H$.

Standard basis for $\mathbb{R}^n$ is the set $\{e_1, \ldots, e_n\}$.

Standard basis for $\mathbb{P}_n$ is the set $\{1, t, t^2, \ldots, t^n\}$.

Th. 4.5 *Let $S = \{v_1, \ldots, v_n\}$ be a set in V and let Span $S = H$.*

1. *If one vector in S is a linear combination of the other vectors in S, then the set formed by removing this vector still spans H.*
2. *If $H \neq \{0\}$, some subset of S is a basis for H*

Basis for Null Space is easy, the null space is its own basis.

Rowops Elementary row operations on a matrix do not affect the linear dependence relations among the columns of the matrix.

Proof of Th. 4.6

Two views a basis can be

1. linearly independent set (as large as possible)
2. spanning set (as small as possible)

Th. 4.6 *The pivot columns of a matrix A form a basis for Col A.*

4.4 Coordinate Systems

Th. 4.7 (Unique Representation Theorem) *Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then for each $x \in V$, there exist unique scalars $[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ such that*

$$x = c_1 b_1 + \cdots + c_n b_n$$

Th. 4.8 (Change of coord is 1-to-1, linear) *Let $\mathcal{B}$ be a basis for a vector space V. Then the coordinate mapping $x \mapsto [x]_{\mathcal{B}}$ is a 1-to-1 linear transformation from V onto $\mathbb{R}^n$*

Coordinates of x relative to the basis $\mathcal{B}$ are the weights $c_1, \ldots, c_n$.

Change of coord.s matrix $P_{\mathcal{B}} = [b_1 \cdots b_n]$.

Change of coordinates in $\mathbb{R}^n$. If there is another basis for $\mathbb{R}^n$, then $x = P_{\mathcal{B}}[x]_{\mathcal{B}}$.

Isomorphism between $\mathbb{R}^n$ and any n dimensional subspace is defined by the transform $x \mapsto [x]_{\mathcal{B}}$.

4.5 Dimensions of a Vector Space

Th. 4.9 (Dimension and Dependence) *If a vector space V has a basis with n vectors, then any set in V containing more than n vectors must be linearly dependent.*

Th. 4.10 (Size of Every Basis) *If a vector space V has a basis of n vectors, then every basis of V has n vectors.*

Finite dimensional vector spaces are spanned by a finite number of vectors. If it's not, it is **infinite dimensional**.

Zero vector space $\{0\}$ is defined to be zero

Example $\dim = \infty$, the set of all polynomials

Dimension of null space is number of free variables

Dimension of column space is number of basic variables

Th. 4.11 *Let H be a subspace of a finite dimensional vector space V. Any linearly independent set can be expanded (add more linearly independent vectors) to form a basis for H.*

Also, H is finite dimensional and:

$$\dim H \leq \dim V$$

Th. 4.12 (Basis Theorem) *Let V be a set with p dimensions. Any set of p linearly independent vectors forms a basis for V.*

Any set of p vectors that span V form a basis for V.

4.6 Rank

Th. 4.13 *Two row equivalent matrices A and B share the same row space.*

If B is in echelon form, the nonzero rows of B form a basis for the row space of A and for B.

Row space set of all linear combinations of the row vectors of a matrix

Th. 4.13

Basis for Col space of a matrix: pivot columns of the matrix

Basis for Row space of a matrix: rows a row equivalent matrix in echelon form

Basis for Nul space of a matrix: vectors in the vector form of the general solution to $Ax = 0$.

Rank of $A = \dim$ Col A.

Dimension of Row space $= \dim$ Col $A = \dim$ Col A^T

Th. 4.14 (The Rank Theorem) *The dimensions of the column space and the row space of an $m \times n$ matrix A are the same. This common dimension is equal to the number of pivot positions in A and*

$$\textit{rank } A + \dim \textit{Nul } A = n$$

Proof of Th. 4.14

Example if A is a 7×9 matrix, what's is rank?

The equation $r + 2 = 9$ must be satisfied, so the rank is 7.

Example Could a 6×9 matrix have $\dim$ Nul $A = 2$?

No. If it did, then the rank should be 7, but there's only 6 columns!

Proof of IMT cont. 4.15

Th. 4.15 (Rank and IMT, cont.) *Let A be an $n \times n$ matrix, these statements are equivalent to the statement that A is an invertible matrix:*

1. *The columns of A form a basis for $\mathbb{R}^n$*
2. *Col $A = \mathbb{R}^n$.*
3. *$\dim$ Col $A = n$*
4. *rank $A = n$*
5. *Nul $A = \{0\}$*
6. *$\dim$ Nul $A = 0$*

5 Eigenvalues and Eigenvectors

5.1 Definition

Eigenvector of an $n \times n$ matrix A is any nonzero vector x such that $Ax = \lambda x$ for some scalar called an **eigenvalue** λ.

Eigenvalue λ exists if there is a nontrivial solution x of $Ax = \lambda x$. This x is an *eigenvector corresponding to* λ. This scalar may be 0.

Find eigenspace given eigenvalue λ: solution set to $Ax = \lambda x$ or equivalently $(A - \lambda I)x = 0$. The null space of the matrix $A - \lambda I$.

Difference Equations can be solved using eigenvalues. If $x_{k+1} = Ax_k$ for $k \in \mathbb{Z}_{+,0}$, then an explicit solution to the system is $x_k = \lambda^k x_0$ for $k \in \mathbb{Z}_+$.

Inductive step this works because: $Ax_k = A(\lambda^k x_0) = \lambda^k(Ax_0) = \lambda^k(\lambda x_0) = x_{k+1}$. Linear combinations of the form $\lambda^k v$ are also solutions.

Th. 5.1 (Eigenvalues of Triangular) *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Th. 5.2 (Linear Independence of Eigenvectors) *If $v_1, \dots, v_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_r$ of a matrix, then the set $\{v_1, \dots, v_r\}$ is linearly independent.*

5.2 Characteristic Equation

Motivation the determinant of a matrix is zero iff it is not invertible; if this is the case then its Null space has non-zero dimension.

Determinants see chapter 3

Characteristic Eqn. a scalar λ is an eigenvalue of an $n \times n$ matrix A iff λ satisfies the characteristic equation: $\det A - \lambda I = 0$. Also called characteristic polynomial of degree n.

Multiplicity of an eigenvalue is its multiplicity in the characteristic equation. Example: the characteristic polynomial $(\lambda - 3)^5 = 0$ reveals an eigenvalue 3 with multiplicity 5.

Similarity $n \times n$ matrices A and B are similar if there exists an invertible matrix P such that $A = PBP^{-1}$.

Similarity Transformation the act of changing A into $P^{-1}AP$.

Application to Dynamical Systems. Let A be an $n \times n$ matrix. Let x_0 be the initial state with coordinates $c_1, \dots, c_n$ with respect to the basis formed by eigenvectors $v_1, \dots, v_n$.

If all the eigenvalues are in $[-1, 1]$, then the system converges to a steady state as $k \to \infty$. It could also diverge, or just converge to zero.

The system $x_{k+1} = Ax_k$ has an explicit solution $x_k = c_1\lambda_1^k v_{\lambda 1} + \dots + c_n \lambda_n^k v_{\lambda n}$.

Th. 5.3 (Invertible Matrix Th. (cont.)) *Let a be an $n \times n$ matrix. Then A is invertible iff*

1. *The number 0 is not an eigenvalue of A*
2. *$\det A \neq 0$*

Th. 5.4 (Properties of Determinants) *— see section 3 —*

Th. 5.5 (Similarity and Eigenvalues) *If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and the same eigenvalues with the same multiplicities.*

WARNING: similarity is not the same as row equivalence. Row operations on a matrix usually change its Eigenvalues.

5.3 Diagonalization

Diagonalizable square matrix is similar to a very specific diagonal matrix.

Eigenvector Basis is formed by the set of all eigenvectors for all eigenvalues of a diagonalizable $n \times n$ matrix. Because it contains enough independent vectors, this set forms a basis for $\mathbb{R}^n$.

Diagonalizing Matrices To factor an $n \times n$ matrix A into form $A = PBP^{-1}$, four steps are required.

1. Find eigenvalues of A. Use computer or characteristic equation.
2. Find n linearly independent vectors of A. Procedure: find the null space of matrix $A - \lambda I$. **Check for linear independence** amongst the vectors.
3. Construct P from the eigenvectors. Order unimportant. **Check P is invertible**.
4. Construct D from the corresponding eigenvalues. **Same order** as P.

Th. 5.6 (Diagonalization Theorem) *If an $n \times n$ matrix A has n distinct eigenvalues, then it is diagonalizable.*

Powers of a matrix with form $A = PDP^{-1}$ can be easily calculated $A^k = PD^kP^{-1}$.

Functions on a matrix with form $A = PDP^{-1}$ can also be easily calculated. For example, to take the sine simply take the sine of the diagonal entries of D: $\sin A = P(\sin D)P^{-1}$. Found by using power series of $\sin, \exp$, etc.

Non-distinct Eigenvalues If the matrix A has less than n eigenvalues, it is still possible for A to be diagonalizable. See Th. 5.7.

Th. 5.7 ($n \times n$ and p distinct eigenvalues) *Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.*

1. *For $1 \le k \le p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.*
2. *The matrix A is diagonalizable iff the sum of the dimensions of the distinct eigenspaces equals n.*
   This only happens when the dimensions of the eigenspace for each eigenvalue λ_k equals the multiplicity of λ_k.
3. *If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an* **eigenvector basis** *for $\mathbb{R}^n$.*

5.4 Eigenvectors and Linear Transformations

Linear Trans. and D a linear transformation $T(x) = Ax$ with a diagonalizable standard matrix A can be represented using a simpler transformation of the form: $u \mapsto Du$.

Linear Transformation Let V, W be vector spaces where $\dim V = n$ and $\dim W = m$. Let $T : V \to W$. Choose two ordered bases $\mathcal{B} = \{b_1, \ldots, b_n\}, \mathcal{C}$ for V, W respectively.

Matrix-based representation for the linear transformation T is $T(v) = [M[x]_{\mathcal{B}}]_{\mathcal{C}}^{-1}$. As far as coordinate vectors go, T is just a left-multiplication by M.

Matrix Relative to Bases $\mathcal{B}$ and $\mathcal{C}$ is the $m \times n$ matrix: $M = \big[[T(b_1)]_{\mathcal{C}} \quad \cdots \quad [T(b_n)]_{\mathcal{C}}\big]$.

Change-of-coordinates matrix is created when $V = W$ and T is the identity function.

Example suppose $\mathcal{B} = \{b_1, b_2\}$ is a basis for V and $\mathcal{C} = \{c_1, c_2, c_3\}$ is a basis for W. Let $T : V \to W$ be a linear transformation with the property that $T(b_1) = 3c_1 - 2c_2 + 5c_3$ and $T(b_2) = 4c_1 + 7c_2 - c_3$.
The matrix M relative to $\mathcal{B}$ and $\mathcal{C}$ is

$$M = [[T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}}] = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

From V into V (linear transformations) If $T : V \to V$, then the matrix M is called the matrix for T relative to $\mathcal{B}$ or $\mathcal{B}$-matrix for T denoted by $[T]_{\mathcal{B}}$. Then $[T(x)]_{\mathcal{B}} = [T]_{\mathcal{B}}[x]_{\mathcal{B}}$.

Example mapping $T : \mathbb{P}_2 \to \mathbb{P}_2$ is defined by $T(a_0 + a_1t + a_2t^2) = a_1 + 2a_2t$. The basis $\mathcal{B} = \{1, t, t^2\}$.
The $\mathcal{B}$-matrix for T is $T(1) = 0; T(t) = 1; T(t^2) = 2t$. The coordinates of these polynomials are used (in order) as columns:

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

For a general polynomial $T(a_0 + a_1t + a_2t^2) = [T]_{\mathcal{B}} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$.

Linear Trans. on $\mathbb{R}^n$ if the transformation is $T(x) = Ax$ and A is diagonalizable, then the $\mathcal{B}$-matrix for T is a diagonal matrix.

Similarity of Matrix Representations

Efficient computation of a $\mathcal{B}$-matrix $P^{-1}AP$ ...

Th. 5.8 (Diagonal Matrix Representation) *Suppose $A = PDP^{-1}$ where D is an $n \times n$ diagonal matrix. If $\mathcal{B}$ is the basis of $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix representation for the transformation $x \mapsto Ax$.*

That is to say, the mappings $x \mapsto Ax$ and $u \mapsto Du$ are the same linear transformation relative to different bases.

5.5 Complex Eigenvalues

Complex eigenvalues allowed? Then A must act on vectors in $\mathbb{C}^n$. Linear algebra can be rebuilt using C^n instead of $\mathbb{R}^n$, but that is not our concern here.

Finding complex eigenvalues , these arise naturally from the Fundamental Theorem of Algebra.

Finding complex eigenvectors use a machine for row reducing complex matrices.

Rotations a real matrix with complex eigenvalues always defines a rotation of vectors.

Complex conjugate of a complex number $\overline{a+bi} = a - bi$. For a matrix, just take the complex conjugate of each entry. Properties for complex numbers, vectors and matrices:

$$\overline{rx} = \overline{r}\overline{x}, \quad \overline{Bx} = \overline{B}\overline{x}, \quad \overline{BC} = \overline{B}\overline{C}, \quad \overline{rB} = \overline{r}\overline{B}$$

Complex to real Let $c \in \mathbb{C}$. Then $\overline{c} \in C$ but $c\overline{c} \in \mathbb{R}$. This is why a real matrix' eigenvalues come in conjugate pairs.

Conjugate pairs if a real matrix A has complex eigenvalue $a + bi$ then $a - bi$ is also an eigenvalue.

Scale and rotate consider the matrix $C = \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$. Its eigenvalues are $\lambda_{1,2} = a \pm bi$. Let $r = |\lambda_1|$ and ϕ be the angle between ray from 0 to (a, b). Then the matrix represents a scaling followed by a rotation transformation: $C = \left[\begin{smallmatrix} r & 0 \\ 0 & r \end{smallmatrix}\right] \left[\begin{smallmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{smallmatrix}\right]$.

Rotation in any matrix with complex eigenvalues A. Any matrix with complex eigenvalues is similar to C. Therefore, multiplication by A represents a change of coordinates wrapping a rotation PCP^{-1}. occur in complex conjugate pairs. Then $A^k x$ can trace out an ellipse instead of a circle.

Th. 5.9 (Complex Eigenvalues and Rotations) *Let A be a real 2×2 matrix with a complex eigenvalue pair $a \pm bi$. where $b \neq 0$ and an associated vector $V \in \mathbb{C}^2$. Then $A = PCP^{-1}$ where $P = [\text{Re}\ v \quad \text{Im}\ v]$ and $C = \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$.*

5.6 Discrete Dynamical Systems

Eigenvector decomposition of a discrete dynamical system $x_{k+1} = Ax_k$ determines the limit of the sequence $\{x_k\}$. If A has m eigenvalues λ_i and x_0 has coordinates c_i with respect to eigenvector basis v_i formed by the eigenvalues then $x_k = c_1\lambda_1^k v_1 + \cdots + c_m\lambda_m^k v_m$.

Convergence/divergence the system converges if every eigenvalue has magnitude less than or equal to 1. The system diverges if at least one eigenvalue is greater than 1 in magnitude.

Trajectory graphing the sequence $x_0, x_1, \ldots$.

2D Systems have several types of trajectories.

Boring solutions are simple when $c = 0$ or when the eigenvalues are 1.

Attractor trajectories converge to the origin if both eigenvalues are real and < 1 in magnitude

Repellor trajectories diverge if both eigenvalues are real and > 1 in magnitude

Saddle point trajectories converge to a line but diverge away from the origin if both eigenvalues are real and one is < 1 and the other is > 1 in magnitude.

Nondiagonal case

Complex eigenvalues are conjugate pairs, so they have the same magnitude.

Spiral outward trajectories diverge and rotate if complex eigenvalues have magnitude > 1.

Spiral inward trajectories converge and rotate if complex eigenvalues have magnitude < 1.

Elliptical trajectory trajectories are rotated if complex eigenvalues have magnitude 1.

5.7 Applications to Differential Equations

Differential Equation

Fundamental set of solutions

Initial value problem

Eigenfunctions

Trajectory

Attractor/sink

Repellor/source

Saddle point

Complex eigenvalues

Spiral point

5.8 Iterative Estimates for Eigenvalues

Strictly Dominant Eigenvalue

The power method can approximate the strictly dominant eigenvalue of a matrix:

1.
2.
3.

Inverse power method provides approximation for any eigenvalue:

1.
2.
3.
4.

6 Orthogonality and Least Squares

6.1 Inner Product, Length, Orthogonality

Inner Product u, v are $n \times 1$ matrices, their inner product is $u^T v$, a scalar. AKA **dot Product**.

Length or norm of a vector defined by $\|v\| = \sqrt{v \cdot v}$. Property: $\|cv\| = c\|v\|$.

Unit Vector has norm 1. **Normalizing** a vector is done by dividing by its norm $u = v\|v\|$, then $\|u\| = 1$.

Distance between vectors u and v is given by $\|u - v\|$.

Orthogonal vectors two vectors are orthogonal, or perpendicular if $u \cdot v = 0$. This is equivalent to having a $90°$ angle between them.

Vector $\perp$ subspace if the vector is perpendicular to every vector in the subspace. Equivalently, it is perpendicular to every basis vector of the subspace.

Orthogonal complement let W be a subspace of $\mathbb{R}^n$. The orthogonal complement of W is $W^\perp$ pronounced "W perp." It is defined as the set of all vectors perpendicular to the vectors in W, $W^\perp = \{v \perp W | v \in \mathbb{R}^n\}$.

New Subspace If W is a subspace of $\mathbb{R}^n$, then $W^\perp$ is a subspace of $\mathbb{R}^n$.

Angle between two vectors u and v is θ and found using: $\cos\theta = u \cdot v / \|u\|\|v\|$. This formula works for higher dimensions, in statistics $\cos\theta$ is called the correlation coefficient between two vectors.

Th. 6.1 (Inner Product Algebra) *Let $u, v, w \in \mathbb{R}^n$ and let $c \in \mathbb{R}$. Then*

1. $u \cdot v = v \cdot u$
2. $(u + v) \cdot w = u \cdot w + v \cdot w$
3. $(cu) \cdot v = u \cdot (cv) = c(u \cdot v)$
4. $u \cdot u \geq 0$ *and* $u \cdot u = 0$ *iff* $u = 0$

Th. 6.2 (Pythagorean Theorem) *Two vectors u and v are orthogonal iff $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.*

Th. 6.3 (Orthog. Complement of a Row Space) *Let A be an $m \times n$ matrix. Then*

$$(\textit{Col } A)^\perp = \textit{Nul } A^T \qquad (\textit{Col } A^T)^\perp = \textit{Nul } A$$

6.2 Orthogonal Sets

Th. 6.4 (Orthog. Set, Linearly Independent) *If S is an orthogonal set of nonzero vectors in $\mathbb{R}^n$ then S is linearly independent and a basis for the subspace spanned by S.*

Orthogonal Set $\{u_1, \ldots, u_p\}$ is formed if each element is perpendicular to every other element. That is, $u_i \cdot u_j = 0$ if $i \neq j$.

Orthogonal Basis A basis for a subspace of $\mathbb{R}^n$ that is also an orthogonal set.

Orthogonal Projection given a vector u it is possible to decompose a vector y into two components: $y = \hat{y} + z$ where $\hat{y} = \alpha u$ and $z \perp u$. The vector $\hat{y}$ is found in terms of u, $\hat{y} = \dfrac{y \cdot u}{u \cdot u} u$. The vector z is easily found, it may be the zero vector.

Geometric Interpretation of Th 6.5 Theorem 6.5 decomposes a vector into a sum of orthogonal projections onto one-dimensional subspaces (lines). Like graph paper with a rectangular grid.

Orthonormal Sets an orthogonal set composed of unit vectors. Very important.

Orthonormal basis an orthogonal set forms such a basis for the subspace spanned by the set.

Orthogonal Matrix a matrix with columns that form an orthonormal set. Notice not an orthogonal set (not as useful).

Rows of an orthogonal matrix also form an orthonormal set.

Meaning of Th 6.7 multiplication by the matrix U preserves length and angle of vectors.

Th. 6.5 (Orthog. Basis and Linear Combo.) *Let $\{u_1, \ldots, u_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then the weights for each $y \in W$ in the linear combination $y = c_1 u_1 + \cdots + c_p u_p$ are given by*

$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

Th. 6.6 (Orthogonal Columns) *An $m \times n$ matrix U has orthonormal columns iff $U^T U = I$.*

Th. 6.7 (Algebra of Orthonormal Columns) *Let U be an $m \times n$ matrix with orthonormal columns and let $x, y \in \mathbb{R}^n$. Then*

1. $\|Ux\| = \|x\|$
2. $(Ux) \cdot (Uy) = x \cdot y$
3. $(Ux) \cdot (Uy) = 0$ *iff* $x \cdot y = 0$

6.3 Orthogonal Projections

Th. 6.8 (Orthogonal Decomposition Theorem) *Let W be a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{u_1, \ldots, u_p\}$. Then each $y \in \mathbb{R}^n$ can be written uniquely as*

$$y = \hat{y} + z$$

where $\hat{y} \in W$, $z \in W^\perp$, $z = y - \hat{y}$, and

$$\hat{y} = \frac{y^T u_1}{u_1^T u_1} u_1 + \cdots + \frac{y^T u_p}{u_p^T u_p} u_p$$

Vector decomposition a vector can be decomposed into a sum of vectors, one in W and one in $W^\perp$.

Orthogonal Projection of y onto W: the vector $\hat{y}$ often written as $\text{proj}_W y$.

Geometric Interpretation of Orthogonal Projections: an orthogonal projection of a vector is the sum of its projectins onto mutually orthogonal one-dimensional subspaces.

Identity If $y \in W$ then $y = \text{proj}_W y$.

Orthogonal matrix has **orthonormal** columns

Th. 6.9 (The Best Approximation) *Let W be a subspace of $\mathbb{R}^n$, $y \in \mathbb{R}^n$ and $\hat{y} = \text{proj}_W y$. Then $\hat{y}$ is the closest point in W to y; in the sense that for all $v \in \mathbb{R}^n$, $v \neq y$:*

$$\|y - \hat{y}\| < \|y - v\|$$

Th. 6.10 (Orthonormal Basis and Projections) *If $\{u_1, \ldots, u_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then*

$$\text{proj}_W y = (y^T u_1) u_1 + \cdots + (y^T u_p) u_p$$

If $U = [u_1 \quad \cdots \quad u_p]$, then for all $y \in \mathbb{R}^n$:

$$\text{proj}_W y = U U^T y$$

6.4 Gram-Schmidt Process

Th. 6.11 (The Gram-Schmidt Process) *Let W be a subspace of $\mathbb{R}^n$ and assume it has a basis $\{x_1, \ldots, x_p\}$. Let*

$$\begin{aligned} v_1 &= x_1 \\ v_2 &= x_2 - \frac{x_2^T v_1}{v_1^T v_1} v_1 \\ &\vdots \\ v_p &= x_p - \frac{x_p^T v_1}{v_1^T v_1} v_1 - \cdots - \frac{x_p^T v_{p-1}}{v_{p-1}^T v_{p-1}} v_{p-1} \end{aligned}$$

The set $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W and $\text{Span}\{v_1, \ldots, v_p\} = \text{Span}\{x_1, \ldots, x_p\} = W$.

A Simple Algorithm for producing an orthogonal basis for a nonzero subspace of $\mathbb{R}^n$. Take two linearly independent vectors, subtract the projection $z = v_2 - \text{proj}_{v_1} v_2$ to get another linearly independent but now orthogonal vector.

Orthonormal Bases an orthogonal basis composed of normalized vectors.

QR Factorization of Matrices finding an orthogonal basis for an $m \times n$ matrix using the Gram-Schmidt process is similar to factoring $A = QR$ where Q is orthogonal and R is square and upper triangular.

Finding R because Q is orthogonal $R = Q^{-1} A = Q^T A$.

Th. 6.12 (QR Factorization) *If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$ where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.*

6.5 Least-Squares Problems

Inconsistent systems arise often in applications. An approximate solution is often demanded.

Least-squares solution of $Ax = b$ is an $\hat{x} \in \mathbb{R}^n$ such that $\|b - A\hat{x}\| \leq \|b - Ax\|$ for all $x \in \mathbb{R}^n$. Note $b \in \mathbb{R}^m$ and A is an $m \times n$ matrix.

Least squares error $\|b - \hat{b}\|$

Soln. of General Least-Squares Problem. Let $Ax = b$ be an inconsistent system. Let $\hat{b} = \text{proj}_{\text{Col } A} b$. Because $\hat{b}$ is in the column space of A, there exist a solution $\hat{x}$ to the equation $A\hat{x} = \hat{b}$.

Optimal This point $\hat{b}$ is the closest point possible to b.

Proof of Th. 6.13 By Th 6.8 $b - A\hat{x}$ is in $(\text{Col } A)^\perp$. Therefore $a_j^T (b - A\hat{x}) = 0$ for every column a_j of A. Since a_j^T is a column of A^T, $A^T b - A^T A\hat{x} = 0$ and $A^T b = A^T A\hat{x}$. $\square$ Conversely, suppose $\hat{x}$ satisfies $A^T A\hat{x} = A^t b$.

Normal Equations of the system $Ax = b$ is this: $A^T Ax = A^T b$.

QR Factorization of A can give more reliable results when using a machine.

Th. 6.13 (Least-Squares and Normal Equations) *The set of least-squares solutions of $Ax = b$ coincides with the nonempty set of solutions of the normal equations $A^T Ax = A^T b$.*

Th. 6.14 (Only 1 Least-Squares Soln.) *The matrix $A^T A$ is invertible iff the columns of A are linearly independent. In this case, there exists only one least squares solution $\hat{x}$:*

$$\hat{x} = (A^T A)^{-1} A^T b$$

Th. 6.15 (Factorization and Linear Indep.) *Given an $m \times n$ matrix A with linearly idnependent columns, let $A = QR$ be a QR factorization of A, Then for each $b \in \mathbb{R}^m$, the unique least-squares solution to $Ax = b$ is*

$$\hat{x} = R^{-1} Q^T b$$

6.6 Applications to Linear Models

Statistical analysis can be made using the least-squares solution of a given system. This is helpful in finding a function that fits a set of data.

Names Design matrix X, parameter vector β, observation vector y.

Least-Squares Lines the line $y = \beta_0 + \beta_1 x$ that minimizes the sum of the squares of the residuals (offsets from the line). The closest fitting line to a set of points.

Regression Coefficients found by finding the least-squares solution β of the system $X\beta = y$.

Example find equation that best fits the points $(2,1),(5,2),(7,3),(8,3)$.

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

The normal equations $X^T X\beta = X^T y$ have solution $\beta_0 = 2/7$ and $\beta_1 = 5/14$. Therefore the least-squares line fitting the data is $y = 2/7 + 5x/14$.

Mean-deviation form compute average $\overline{x}$ of the original x-values and form a new variable $x^* = x - \overline{x}$. Then the X matrix is orthogonal and its QR factorization can be used for a simpler computation of the solution.

Residual Vector $\epsilon = y - X\beta$.

General Linear Model any equation of the form $y = X\beta + \epsilon$. The matrix X need not model a linear equation, it could be an entirely different curve for the data.

Fitting of Other Curves if it is best to fit the data to a general function $y = \beta_0 f_0(x) + \cdots + \beta_k f_k(x)$ then the design matrix X is simply changed to

$$X = \begin{bmatrix} f_0(x_1) & \cdots & f_k(x_k) \\ \vdots & \ddots & \vdots \\ f_0(x_n) & \cdots & f_k(x_n) \end{bmatrix}$$

Multiple Regression if an experiment involves two independent variables u, v and one dependent variable, we can find a **trend surface**. Data: $(u_0, v_0, y_0), \ldots, (u_n, v_n, y_n)$. Function: $y = \beta_0 f_0(u,v) + \cdots + \beta_k f_k(u,v)$. The design matrix is computed as above.

6.7 Inner Product Spaces

Inner Product on a Vector Space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$. It must satisfy the following axioms for all $u, v, w \in V$ and all $c \in \mathbb{R}$:

1. $\langle u, v \rangle = \langle v, u \rangle$
2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
3. $\langle cu, v \rangle = c\langle u, v \rangle$
4. $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = 0$

Inner Product Space is a vector space with an inner product

Length $\|v\| = \sqrt{\langle v, v \rangle}$ for all $v \in V$.

Distance the distance between $v, u \in V$ is $\|u, v\|$.

Orthogonality two vectors are orthogonal if $\langle u, v \rangle = 0$.

Projection of a vector onto a subspace with an orthogonal basis can be constructed as usual. Has the properties described in the Orthognal Decomposition Theorem and the Best Approximation Theorem.

Gram-Schmidt Process works as usual. Can be proven.

Best Approximation in Inner Product Spaces.

Inner Product for $C[a,b]$

Th. 6.16 (Cauchy-Schwarz Inequality) *For all $u, v \in V$:*

$$|\langle u, v \rangle| \leq \|u\|\|v\|$$

Th. 6.17 (Triangle Inequality) *For all $u, v \in V$:*

$$\|u + v\| \leq \|u\| + \|v\|$$

6.8 Applications of Inner Product Spaces

Weighted least-squares

Trend Analysis of Data

Fourier Series