Introduction to Linear Algebra Notes

Nathan Ueda

March 6, 2024

Contents

1 Vectors and Matrices

1.1 Vectors and Linear Combinations

Vector Length: For a vector $\boldsymbol{v} \in \mathbb{R}^n$, its length is:

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

In words, the length of a vector is the square root of the sum of the squared components.

Given two vectors in $\mathbb{R}^2$ $\boldsymbol{v}, \boldsymbol{w}$ with their tail starting from the origin

- If they lie on the same line, the vectors are *linearly dependent*.
- If they do not lie on the same line, the vectors are *linearly independent*.

Therefore, the combinations $c\boldsymbol{v} + d\boldsymbol{w}$ fill the $x - y$ plane unless $\boldsymbol{v}$ is in line with $\boldsymbol{w}$.

To fill m-dimensional space, we need m independent vectors, with each vector having m components.

1.2 Lengths and Angles from Dot Products

Dot Product: For two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$, their dot product is:

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_1 w_1 + \cdots + v_n w_n$$

The dot product of two vectors tells us what amount of one vector goes in the direction of another. It tells us how much these vectors are working together.

- $\boldsymbol{v} \cdot \boldsymbol{w} > 0$: The vectors point in somewhat similar directions. In other words, the angle between the two vectors is less than 90 degrees.
- $\boldsymbol{v} \cdot \boldsymbol{w} = 0$: The vectors are perpendicular. In other words, the angle between the two vectors is 90 degrees.
- $\boldsymbol{v} \cdot \boldsymbol{w} < 0$: The vectors point in somewhat opposing directions. In other words, the angle between the two vectors is greater than 90 degrees.

Dot Product Rules (for two vectors, $\boldsymbol{v}, \boldsymbol{w}$):

- $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$
- $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$
- $(c\boldsymbol{v}) \cdot \boldsymbol{w} = c(\boldsymbol{v} \cdot \boldsymbol{w})$

Cosine Formula: If $\boldsymbol{v}$ and $\boldsymbol{w}$ are nonzero vectors, then:

$$\cos\theta = \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|}$$

Unit Vectors: A vector is a unit vector if its length is 1.
For a vector $\boldsymbol{u} \in \mathbb{R}^n$:

$$\|\boldsymbol{u}\| = 1$$

For any vector $\boldsymbol{v} \in \mathbb{R}^n$, as long as $\boldsymbol{v} \neq 0$, dividing $\boldsymbol{v}$ by its length will result in a unit vector. In other words:

$$\boldsymbol{u} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$$

Cauchy-Schwarz Inequality:

$$|\boldsymbol{v} \cdot \boldsymbol{w}| \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|$$

In words, the absolute value of the dot product of two vectors is no greater than the product of their lengths.

Triangle Inequality:

$$\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$$

In words, the length of any one side (in this case $\|\boldsymbol{v} + \boldsymbol{w}\|$) of a triangle is at most the sum of the length of the other triangle sides.

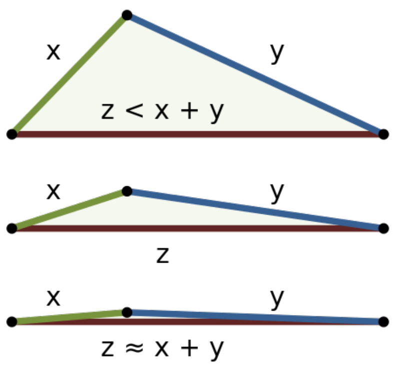

Figure 1: This Squeeze Theorem

1.3 Matrices and Their Column Spaces

Independence: Columns are independent when each new column is a vector that we don't already have as a combination of previous columns. The only combindation of columns that produces $A\boldsymbol{x} = (0, 0, 0)$ is $\boldsymbol{x} = (0, 0, 0)$.

Column Space: The column space, $\mathbf{C}(A)$, contains all vectors $A\boldsymbol{x}$. In other words, it contains all combinations of the columns.
The **span** of the columns of A is the column space.

Rank: The number of independent columns of a matrix. This is equivalent to saying the rank is the number of pivots in a matrix.

Rank Rules:

- $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$

1.4 Matrix Multiplication AB and CR

To multiply two matrices AB, take the dot product of each row of A with each column of B. The number in row i, column j of AB is (row i of A) $\cdot$ (column j or B).

When $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$:

- $AB \in \mathbb{R}^{m \times p}$
- mp dot products are needed to carry out the matrix multiplication (one for each entry in the matrix AB).

Matrix Multiplication Rules:

- Associative: $(AB)C$
- Distributive: $A(B + C) = AB + BC$
- Not Commutative: In general $AB \neq BA$

2 Solving Linear Equations Ax = b

2.1 Elimination and Back Substitution

For a matrix $A \in \mathbb{R}^{n \times n}$, there are three outcomes for $A\boldsymbol{x} = \boldsymbol{b}$:

1. No solution
   - $\boldsymbol{b}$ is not in the column space of A
   - This occurs when the columns of A are dependent and $\boldsymbol{b}$ is not in $C(A)$ ($\cancel{32}$ = any value but 32).

$$\begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} = \begin{bmatrix} 16 \\ \cancel{32} \end{bmatrix}$$

2. Exactly 1 solution
   - A has independent columns and an inverse matrix A^{-1}
3. Infinitely many solutions
   - Columns of A are dependent.
   - This occurs when the columns of A are dependent and $\boldsymbol{b}$ is in $C(A)$

$$\begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} = \begin{bmatrix} 16 \\ 32 \end{bmatrix}$$

Elimination: A system that allows us to determine if $A\boldsymbol{x} = \boldsymbol{b}$ has no solution, 1 solution, or infinitely many solutions. The goal of elimination is to transform A to an upper triangular matrix, U.

Elimination allows us to discover the number of pivots in $A \in \mathbb{R}^{n \times n}$ by creating U. If there are n pivots in U, U has full rank. This implies A has exactly one solution.

Back Substitution: If U has full rank, back substituion allows us to find the solution.

2.2 Elimination Matrices and Inverse Matrices

Elimination

- The basic elimination step subtracts a multiple of ℓ_{ij} of equation j from equation i

Inverse Matrices

- If A is invertible, the one and only solution to $A\boldsymbol{x} = \boldsymbol{b}$ is $\boldsymbol{x} = A^{-1}\boldsymbol{b}$
- Only square matrices can have inverses
- Invertible $\equiv$ non-singular $\equiv$ non-zero determinant $\equiv$ independent columns
- Not invertible $\equiv$ singular $\equiv$ zero determinant $\equiv$ dependent columns
- If a matrix is invertible, the inverse is unique
- A triangular matrix has an inverse so long that it has no zero diagonal entries

2.3 Matrix Computations and A = LU

Elimination without row exchanges factors A into LU. We can't find an LU decomposition if row exchanges are needed during elimination.

Gauss-Jordan elimination:

- An algorithm that allows us to determine if the inverse of a matrix exists, and if it does it exist, it allows us to determine what the inverse is.
- Augment A by I, that is $[A\ I]$, and through elementary row operations, transform this matrix to $\left[I\ A^{-1}\right]$

3 The Four Fundamental Subspaces

3.1 Vector Spaces and Subspaces

To be a vector space means that all linear combinations $c\boldsymbol{v} + d\boldsymbol{w}$ of the vectors or matrices stay inside that space.

Subspaces:

- A subspace is a vector space entirely contained within another vector space
- All linear combinations of vectors in the subspace stay in the subspace
- Every subspace contains the zero vector
- Subspaces of $\mathbb{R}^3$:
  - The single vector $(0, 0, 0)$
  - Any line through $(0, 0, 0)$
  - Any plane through $(0, 0, 0)$

  - The whole space $\mathbb{R}^3$

Column Space ($C(A)$):

- The column space consists of all linear combinations of the columns
- To solve $A\boldsymbol{x} = \boldsymbol{b}$ is to express $\boldsymbol{b}$ as a combination of the columns
- The right side, $\boldsymbol{b}$, has to be in the column space produced by A, or $A\boldsymbol{x} = \boldsymbol{b}$ has no solution

Row Space ($C(A^T)$):

- The row space of A is the column space of A^T
- The rank of $A =$ rank of A^T

Span:

- The span of vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is the set of all of their linear combinations.
- In other words, this tells us, given some set of vectors, which vectors are able to be created by taking a linear combination of the vectors in the set. It is the vector space we can reach (span) by taking linear combindations of the set of vectors.
- Independence is not required by the word span.

3.2 Computing the Nullspace by Elimination: $\mathbf{A = CR}$

Nullspace ($N(A)$):

- Contains all solutions $\boldsymbol{x}$ to $A\boldsymbol{x} = 0$ including $\boldsymbol{x} = 0$
- If A is invertible, then the nullspace contains only the zero vector (no special solutions)
- If A has n columns, r of which are independent, then there are $n - r$ vectors in the nullspace
- Elimination does not change the nullspace (though it does change the column space)
- If $n > m$ (more columns than rows), then there is at least one free variable (can't have n independent columns in $\mathbb{R}^m$). Therefore, there is at least one nonzero solution
- The dimension of the nullspace is the number of free variables in a matrix

Echelon Form (R):

- Echelon form requirements:
  1. All rows having only zeros are at the bottom
  2. The leading entry (pivot) for each nonzero row (leftmost nonzero entry), is on the right of the leading entry of every row above
- The result of Gaussian elimination on any matrix, square or otherwise
- If the matrix is square and invertible, echelon form is an upper triangular matrix, U
- Can be viewed as a generalization of upper triangular form for rectangular matrices

- Example:

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & 1 & a_6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Reduced Row Echelon Form Requirements (R_0):

- In row echlon form
- The leading entry (pivot) of each nonzero row is a 1
- Each column containing a leading 1 has zeros in all entries above the leading 1.
- Example:

$$\begin{bmatrix} 1 & 7 & 0 & 8 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Algorithm (must be in row echelon form):
  1. Divide pivots by themselves to make them all 1.
  2. Zero out all entires above and below the pivots.

  The result will have the identity matrix in the pivot columns (and the remaining columns will be the special columns)

Finding the Nullspace:
Goal: Given some matrix A, find all solutions $A\boldsymbol{x} = 0$

1. Start with some matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}$$

2. Do elimination to echelon form

$$R = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

3. Do elimination from row echelon form to reduced row echelon form to find the pivot columns and free columns

$$R_0 = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

   Takeaways:

   - $r = 2 =$ number of pivots (coming from columns 1 and 3)
   - 2 free variables (coming from columns 2 and 4)
     - Free variables, which in this case are x_2 and x_4 can take on any value
     - There is one special solution for each free variable

- The special solution for each free variable is found by setting that free variable to 1 and the rest of the free variables to 0

4. Find the special solutions for the free variables using back substituion

   Recall our system of linear equations:

$$x_1 + 2x_2 - 2x_4 = 0$$
$$x_3 + 2x_4 = 0$$

   (a) Let $x_2 = 1, x_4 = 0$. Equation 2 gives us $x_3 = 0$ and equation 1 gives us $x_1 = -2$. The result is a vector in the nullspace:

$$s_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

   (b) Let $x_2 = 0, x_4 = 1$. Equation 2 gives us $x_3 = -2$ and equation 1 gives us $x_1 = 2$. The result is a vector in the nullspace:

$$s_2 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

   Takeaways:

   - These vectors form a basis for the nullspace of A.
   - Any multiple of these 2 vectors (special solutions) are in the nullspace.
   - The nullspace contains exactly all the linear combinations of the special solutions.
   - In other words, for any 2 constants c, d, the nullspace is:

$$cs_1 + ds_2$$

3.3 The Complete Solution to Ax = b

$A\boldsymbol{x} = \boldsymbol{b}$ is solvable when $\boldsymbol{b}$ is in $C(A)$

Finding the Complete Solution to $A\boldsymbol{x} = \boldsymbol{b}$:

1. Augment matrix A to account for the non-zero $\boldsymbol{b}$ on the right side

   $A \rightarrow [A\ \boldsymbol{b}]$

2. Reduce $[A\ \boldsymbol{b}] \rightarrow [U\ \boldsymbol{c}]$

$$[A\ \boldsymbol{b}] = \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & -2 & -2 & b_3 - 3b_1 \end{bmatrix} \rightarrow$$
$$\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} = [U\ \boldsymbol{c}]$$

- Columns 1 and 3 will be the pivot columns (x_1, x_3) and columns 2 and 4 will be the free columns (free variables x_2, x_4)
- Condition for solvability: $b_3 + b_2 - 5b_1 = 0$

3. Reduce $[U\ \boldsymbol{c}] \to [R_0\ \boldsymbol{d}]$

$$[U\ \boldsymbol{c}] \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 1 & 1 & \frac{1}{2}b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 4b_1 - \frac{3}{2}b_2 \\ 0 & 0 & 1 & 1 & \frac{1}{2}b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} = [R_0\ \boldsymbol{d}]$$

4. Find a particular solution (there is only one), $\boldsymbol{x}_\text{p}$, for

$$A\boldsymbol{x} = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix}$$

Note that this $\boldsymbol{b}$ was given in the problem

(a) Set all free variables to 0
$x_1 + 2x_2 + 2x_4 = 4b_1 - \frac{3}{2}b_2$
$x_3 + x_4 = \frac{1}{2}b_2 - b_1$

Let $x_2 = 0, x_4 = 0$
$x_1 = 4b_1 - \frac{3}{2}b_2$
$x_3 = \frac{1}{2}b_2 - b_1$

(b) Solve $A\boldsymbol{x} = \boldsymbol{b}$ for the pivot variables

$$\boldsymbol{x}_\text{p} = \begin{bmatrix} 4b_1 - \frac{3}{2}b_2 \\ 0 \\ \frac{1}{2}b_2 - b_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4(0) - \frac{3}{2}(6) \\ 0 \\ \frac{1}{2}(6) - 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -9 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

5. Find the nullspace

Recall our system of linear equations (and set to 0 since we're solving nullspace):

$$x_1 + 2x_2 + 2x_4 = 0$$

$$x_3 + x_4 = 0$$

(a) Let $x_2 = 1, x_4 = 0$. Equation 2 gives us $x_3 = 0$ and equation 1 gives us $x_1 = -2$. The result is a vector in the nullspace:

$$\boldsymbol{s_1} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

(b) Let $x_2 = 0, x_4 = 1$. Equation 2 gives us $x_3 = -1$ and equation 1 gives us $x_1 = -2$. The result is a vector in the nullspace:

$$\boldsymbol{s_2} = \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

For any 2 constants c_1, c_2, the nullspace is:

$$\boldsymbol{x}_\mathrm{n} = c_1 \boldsymbol{s_1} + c_2 \boldsymbol{s_2}$$

6. Sum the particular and nullspace solutions to get the complete solution

$$\boldsymbol{x} = \boldsymbol{x}_\mathrm{p} + \boldsymbol{x}_\mathrm{n}$$

$$\boldsymbol{x} = \begin{bmatrix} -9 \\ 0 \\ 3 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Properties for every matrix A with full column rank ($r = n$) (a pivot in every column):

- All columns of A are pivot columns (independent). No free variables
- The nullspace, $N(A)$, contains only the zero vector $\boldsymbol{x} = 0$
- There are 0 or 1 solutions to $A\boldsymbol{x} = \boldsymbol{b}$

Properties for every matrix A with full row rank ($r = m$) (a pivot in every row):

- All rows of A are pivot rows and R_0 has no zero rows ($R_0 = R$). $n - r$ free variables
- $A\boldsymbol{x} = \boldsymbol{b}$ has a solution for every right side $\boldsymbol{b}$
- The column space of A is the whole space $\mathbb{R}^m$
- If $m < n$ then $A\boldsymbol{x} = \boldsymbol{b}$ is undetermined (many solutions)

Possibilities for Linear Equations:

1. $r = m, r = n$: square and invertible $\rightarrow A\boldsymbol{x} = \boldsymbol{b}$ has 1 solution
2. $r = m, r < n$: short and wide $\rightarrow A\boldsymbol{x} = \boldsymbol{b}$ has ∞ solutions
3. $r < m, r = n$: tall and thin $\rightarrow A\boldsymbol{x} = \boldsymbol{b}$ has 0 or 1 solution
4. $r < m, r < n$: not full rank $\rightarrow A\boldsymbol{x} = \boldsymbol{b}$ has 0 or ∞ solutions

3.4 Independence, Basis, and Dimension

Linear Independence (below are different ways of saying a matrix has independent columns)

- Having independent vectors essentially means there are no extra, unnecessary vectors. In other words, each vector is necessary, as no vector is a linear combination of the other vectors
- Different ways of saying a matrix has independent columns
  - The columns of A are linearly independent when the only solution to $A\boldsymbol{x} = \mathbf{0}$ is $\boldsymbol{x} = \mathbf{0}$
  - The columns of A are linearly independent when there is no combination of the columns of A, except for the zero vector, solve $A\boldsymbol{x} = \mathbf{0}$
  - The columns of A are linearly independent when its nullspace, $N(A)$, contains only the zero vector

Span

- The span of a set of vectors is defined as the set of all linear combinations of The set of vectors
- Saying that vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ span a space means the space consists of all linear combinations of those vectors
- In $\mathbb{R}^n$
  - $n-1$ independent columns, cannot span all of $\mathbb{R}^n$ (that being said, they do still span a space; however, the space they span is just not all of $\mathbb{R}^n$. In fact, they would even form a basis for this space (since the columns are independent))
  - n independent columns are needed to span all of $\mathbb{R}^n$
  - $n+1$ vectors cannot be independent (only n pivots, so even if we have n pivots, one vector will still be dependent)
- The vectors, $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$, may or may not be independent
  - For example, the columns of a matrix span its column space. These column vectors may be dependent

Basis

- In the special case that vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ are independent (meaning we have just the right amount of vectors to describe the space, any less vectors and we wouldn't fully define the space, any more would be redundant), we say these vector form a basis for the space. In other words, a basis is a special case of a span
- More formally, a basis for a space is a sequence of vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ that have 2 properties
  1. the vectors are independent
  2. the vectors span the space
- The vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ are a basis for $\mathbb{R}^n$ exactly when they are the columns of an n by n invertible matrix
- To see if some vectors form a basis we would

   1. Put the vectors in the columns of a matrix
   2. Do elimination
   3. See if all the columns are pivot columns, if so, they form a basis, else, they don't

- All bases (plural for basis) for a vector space have the same number of vectors
  - For example, there are many sets of vectors that form a basis for $\mathbb{R}^n$ and they all have n vectors

Dimension

- The number of vectors in any and every basis is the dimension of the space
- The dimension of the column space equals the rank of the matrix which equals the number of pivot columns

3.5 Dimensions of the Four Subspaces

The Four Fundamental Subspaces

1. Column Space: $C(A^T)$
2. Row Space: $C(A^T)$
3. Nullspace: $N(A)$
4. Left Nullspace: $N(A^T)$

Column Space

- Contains all the combinations of the columns of a matrix, A
- In $\mathbb{R}^m$ since each column has m components
- The basis for the column space is the pivot columns (number of pivot columns is the rank)
- dim $C(A) = r$

Row Space

- Contains all combinations of the rows of a matrix, A. In other words, it contains all combinations of the columns of A^T
- In $\mathbb{R}^n$ since each row has n components
- The basis for the row space is the first r rows of R_0 (A has a different column space than R_0 but A has the same row space as R_0 since we reached R_0 by doing operations on the rows)
- dim $C(A^T) = r$ (same as the dimension of the column space)

Nullspace

- Contains all solutions $\boldsymbol{x}$ to $A\boldsymbol{x} = \mathbf{0}$, including $\boldsymbol{x} = 0$
- In $\mathbb{R}^m$ since each solution $\boldsymbol{x}$ has m components

- The basis for the nullspace is the special solutions for A (for each free variable, we have one special solution)
- $\dim N(A) = n - r$ (number of free variables in A)

Left Nullspace

- Contains all solutions $\boldsymbol{y}$ to $A^T\boldsymbol{y} = \mathbf{0}$, including $\boldsymbol{y} = 0$
- This is called the left nullspace since $A^T\boldsymbol{y} = \mathbf{0}$ is equivalent to $\boldsymbol{y}^T A = \mathbf{0}^T$ where $\boldsymbol{y}$ is acting on the left
- In $\mathbb{R}^n$ since each solution $\boldsymbol{y}$ has n components
- The basis for the left nullspace is the special solutions for A^T (for each free variable, we have one special solution)
- $\dim N(A^T) = m - r$ (number of free variales in A^T)

Misc.

- The row space and null space are in $\mathbb{R}^n$, have dimensions r and $n - r$, and add to n
- The column space and left nullspace are in $\mathbb{R}^m$, have dimensions r and $m - r$, and add to m

4 Orthogonality

4.1 Orthogonality of Vectors and Subspaces

Orthogonality of Vectors

- 2 vectors are orthogonal when they are perpendicular to each other.
- 2 vectors $\boldsymbol{v}, \boldsymbol{w}$ are orthogonal when their dot product is 0, that is, $\boldsymbol{v}^T\boldsymbol{w} = 0$ (since 90 degree angles between 2 vectors have a dot product of 0).
- If 2 vectors $\boldsymbol{v}, \boldsymbol{w}$ are orthogonal, they form a right triangle and the Pythagorean theorem holds, that is, $\|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 = \|\boldsymbol{v} + \boldsymbol{w}\|^2$, where $\boldsymbol{v} + \boldsymbol{w}$ is the hypotenuse.
- The zero vector is orthogonal to every vector.

Orthogonality of Subspaces

- Subspaces V and W are orthogonal when $\boldsymbol{v}^T\boldsymbol{w} = 0$ for every $\boldsymbol{v} \in V$ and every $\boldsymbol{w} \in W$.
- The nullspace of A is orthogonal to the row space of A.
  - For $A\boldsymbol{x} = 0$ recall $\boldsymbol{x}$ denotes vectors in the nullspace.
  - Every row has a zero dot product with $\boldsymbol{x}$, therefore every combination of the rows is perpendicular to $\boldsymbol{x}$.
  - This shows the whole row space is orthogonal to the whole nullspace.
- The left nullspace of A is orthogonal to the column space of A.
  - This can be seen by applying the same logic as above.

- The only vector in 2 orthogonal subspaces is the zero vector.

Orthogonal Complements

- If two orthogonal subspaces account for the whole space (their dimensions add to the whole space), they are orthogonal complements.
- 2 orthogonal lines in $\mathbb{R}^3$ could not be orthogonal complements because their summed dimensions is 2 instead of 3.
- A line orthogonal to a plane in $\mathbb{R}^3$ are orthogonal complements because their summed dimensions is 3.
- The orthogonal complement $V^\perp$ of V contains all vectors orthogonal to V.
- The nullspace (dim $n - r$) is the orthogonal complement of the row space (dim r) (in $\mathbb{R}^n$).
- The left nullspace (dim $m - r$) is the orthogonal complement of the column space (dim r) (in $\mathbb{R}^m$).

4.2 Projections onto Lines and Subspaces

Why Project?

- $A\boldsymbol{x} = \boldsymbol{b}$ may have no solution ($A\boldsymbol{x}$ may not be in the column space of $\boldsymbol{b}$).
- Therefore, the plan is to solve the closest problem we can.
- Clearly, $A\boldsymbol{x}$ is in the column space of A, which, in this case, is not the same as the column space of $\boldsymbol{b}$.
- Therefore, we will change $\boldsymbol{b}$ to the vector that is closest to the column space of A.
- Instead, we solve $A\hat{\boldsymbol{x}} = \boldsymbol{p}$, where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of A.
- Summary: If $A\boldsymbol{x} = \boldsymbol{b}$ has no solution, solve $A\hat{\boldsymbol{x}} = \boldsymbol{p}$, which is the closest problem to $A\boldsymbol{x} = \boldsymbol{b}$ that has a solution. $P\boldsymbol{b}$ projects $\boldsymbol{b}$ into the column space.

Projection on a Line

- Given 2 vectors, $\boldsymbol{a}, \boldsymbol{b}$ as depicted below

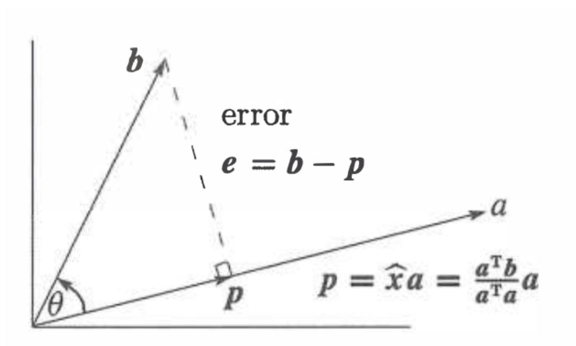

Figure 2: 2 vectors $\boldsymbol{a}, \boldsymbol{b}$

- The projection of $\boldsymbol{b}$ onto $\boldsymbol{a}$ is the closest point from $\boldsymbol{b}$ to $\boldsymbol{a}$ (the dotted line perpendicular to $\boldsymbol{a}$ going toward the tip of $\boldsymbol{b}$).
- The projection $\boldsymbol{p}$ of is the part of $\boldsymbol{b}$ that is along the line.
- The projection $\boldsymbol{p}$ is some multiple of $\boldsymbol{a}$, let's call it $\hat{\boldsymbol{x}}$, that is, $\boldsymbol{p} = \hat{\boldsymbol{x}}\boldsymbol{a}$.
- The error (the part of $\boldsymbol{b}$ that is not part of the projection), which is perpendicular to $\boldsymbol{a}$, is $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \boldsymbol{b} - \hat{\boldsymbol{x}}\boldsymbol{a}$.
- Since $\boldsymbol{a} \perp \boldsymbol{p}, \boldsymbol{a}^T\boldsymbol{p} = \boldsymbol{a}^T(\boldsymbol{b} - \hat{\boldsymbol{x}}\boldsymbol{a}) = 0$.
- Simplifying, we get $\hat{\boldsymbol{x}} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}$.
- Since $\boldsymbol{p} = \hat{\boldsymbol{x}}\boldsymbol{a}, \boldsymbol{p} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}$.
- Doing some slight rearranging, we can see the projection is carried out by a matrix, which we will denote P. In this case, we have $\boldsymbol{p} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a} = \frac{\boldsymbol{a}^T\boldsymbol{a}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{b} = P\boldsymbol{b}$.

Properties of P

- Symmetric: $P = P^T$.
- Its square is itself: $P^2 = P$.

Projection on a Plane

- Given a vector, $\boldsymbol{b}$ and the column space of A, S as depicted below

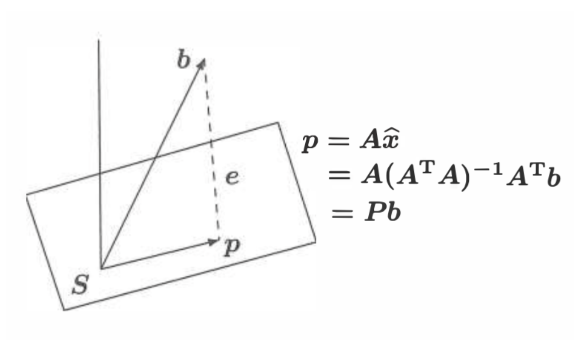

Figure 3: vector $\boldsymbol{a}$, column space of A, S

- Since A is a plane, the matrix of A is made up of 2 independent vectors, that is,

$$A = \begin{bmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} \end{bmatrix}$$

  where $\boldsymbol{a_1}, \boldsymbol{a_2}$ form a basis for A.
- The projection of $\boldsymbol{b}$ onto S is the closest point from $\boldsymbol{b}$ to S (the dotted line perpendicular to S going toward the tip of $\boldsymbol{b}$).
- The projection $\boldsymbol{p}$ of is the part of $\boldsymbol{b}$ that is along the plane.
- The projection $\boldsymbol{p}$ is some multiple of A, let's call it $\hat{\boldsymbol{x}}$, that is, $\boldsymbol{p} = \hat{\boldsymbol{x}}_1\boldsymbol{a_1} + \hat{\boldsymbol{x}}_2\boldsymbol{a_2} = A\hat{\boldsymbol{x}}$.

- The error (the part of $\boldsymbol{b}$ that is not part of the projection), which is perpendicular to S, is $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \boldsymbol{b} - A\hat{\boldsymbol{x}}$.
- Since $\boldsymbol{S} \perp \boldsymbol{p}$

$$\boldsymbol{a}_1^T \boldsymbol{p} = \boldsymbol{a}_1^T(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0$$
$$\boldsymbol{a}_2^T \boldsymbol{p} = \boldsymbol{a}_2^T(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0$$
$$A^T(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0$$
$$A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

- Simplifying, we get $\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$.
- Since $\boldsymbol{p} = A\hat{\boldsymbol{x}}, \boldsymbol{p} = A(A^T A)^{-1} A^T \boldsymbol{b}$.
- Doing some slight rearranging, we can see the projection is carried out by a matrix, which we will denote P. In this case, we have $\boldsymbol{p} = A(A^T A)^{-1} A^T \boldsymbol{b} = P\boldsymbol{b}$.
- This idea easily extends to n vectors.

Special Cases of Projection

1. If $\boldsymbol{b}$ is already in the column space, $P\boldsymbol{b} = \boldsymbol{b}$.
2. If $\boldsymbol{b}$ is perpendicular to the column space, $P\boldsymbol{b} = 0$.

The average case has a b that has a component in the column space and a component perpendicular to column space.

4.3 Least Squares Approximations

- Often we have data points that cannot be fit perfectly with a line.

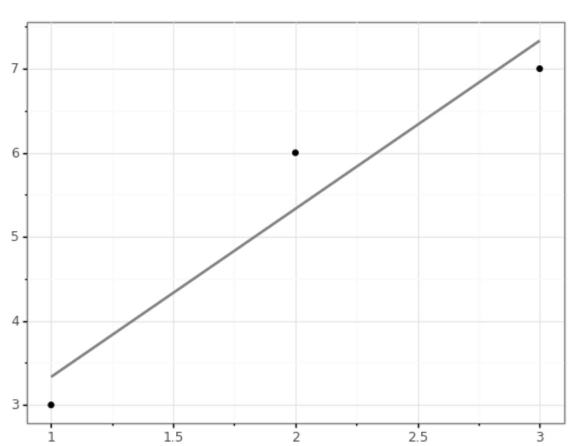

Figure 4: x axis is $\boldsymbol{t}$, y axis is $\boldsymbol{b}$

- The line is defined as

$$\boldsymbol{b} = C + D\boldsymbol{t}$$

- The equations we would like to solve, but can't, are

$$C + 1D = 3$$

$$C + 2D = 6$$

$$C + 3D = 7$$

- Putting these equations into matrix form, we get

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

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$$

  but note this matrix form is still not solvable.

- The line we create is the line that minimizes the overall error, where we define the overall error as the sum of squared error (sum the residuals squared).

- The vector $\boldsymbol{e}$ holds the residuals (the difference in vertical height between the actual point and the line), where $\boldsymbol{e}$ is defined as

$$\boldsymbol{e} = A\boldsymbol{x} - \boldsymbol{b}$$

- Since we want to minimize the sum of squared errors, we want to minimize the length of $\boldsymbol{e}$ squared (the length is the sum the components squared, then take the square root of the result, then we square this result to nullify the square root).

$$\|\boldsymbol{e}\|^2 = \|A\boldsymbol{x} - \boldsymbol{b}\|^2$$

- Though we can't solve

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

  we can solve

$$A\hat{\boldsymbol{x}} = \boldsymbol{p}$$

  where $\boldsymbol{p}$ is the part of $\boldsymbol{b}$ that is in the column space and

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$$

  is the least squares solution (the line that minimizes the sum of square errors).

- The equation for $\hat{\boldsymbol{x}}$ is (Strang said this is the most important equation in statistics)

$$A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 16 \\ 36 \end{bmatrix}$$

- Putting this together, we get the equations

$$3\hat{C} + 6\hat{D} = 16$$

$$6\hat{C} + 14\hat{D} = 36$$

So

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 4/3 \\ 2 \end{bmatrix}$$

- Solving for $\hat{C}$ and $\hat{D}$, we get

$$\hat{D} = 2$$
$$\hat{C} = \frac{4}{3}$$

- So, the best line is

$$\frac{4}{3} + 2t$$

- We can now find the vector $\boldsymbol{p}$ (the part of $\boldsymbol{b}$ in the column space)

$$\boldsymbol{p} = \begin{bmatrix} \frac{4}{3} + 2(1) \\ \frac{4}{3} + 2(2) \\ \frac{4}{3} + 2(3) \end{bmatrix} = \begin{bmatrix} \frac{4}{3} + 2 \\ \frac{4}{3} + 4 \\ \frac{4}{3} + 6 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} + \frac{6}{3} \\ \frac{4}{3} + \frac{12}{3} \\ \frac{4}{3} + \frac{18}{3} \end{bmatrix} = \begin{bmatrix} 10/3 \\ 16/3 \\ 22/3 \end{bmatrix}$$

- We can also calculating the vector $\boldsymbol{e}$

$$\boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} b_1 - p_1 \\ b_2 - p_2 \\ b_3 - p_3 \end{bmatrix} = \begin{bmatrix} 3 - 10/3 \\ 6 - 16/3 \\ 7 - 22/3 \end{bmatrix} = \begin{bmatrix} 9/3 - 10/3 \\ 18/3 - 16/3 \\ 21/3 - 22/3 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix}$$

- Some notes
  - $\boldsymbol{b} = \boldsymbol{p} + \boldsymbol{e}$
  - $\boldsymbol{b} \perp \boldsymbol{e}$
- Fitting this line is linear regression.
- If A has independent columns, then $A^T A$ is invertible.

4.4 Orthonormal Bases and Gram-Schmidt

Orthonormal Basics

- Orthogonal: Vectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are orthogonal when their dot products $\boldsymbol{q}_i \cdot \boldsymbol{q}_j = 0$ whenever $i \neq j$.
- Normal: A vector $\boldsymbol{u}$ is normal when its length is one, that is, $\|u\| = 1$.
- Orthonormal: $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are orthonormal when they are both orthogonal and normal.

- More formally, the n vectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are orthonormal if

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0 & \text{when } i \neq j \quad \text{orthogonal vectors} \\ 1 & \text{when } i = j \quad \text{unit vectors } \|\boldsymbol{q}_i\| = 1 \end{cases}$$

Orthonomal Martix

- A matrix Q is orthonormal if it has orthonormal columns.

$$Q^T Q = \begin{bmatrix} -\boldsymbol{q}_1^T- \\ -\boldsymbol{q}_2^T- \\ -\boldsymbol{q}_n^T- \end{bmatrix} \begin{bmatrix} | & | & | \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_n \\ | & | & | \end{bmatrix} = I$$

- If Q is square, then $Q^T Q = I$ means that $Q^T = Q^{-1}$.

Why Orthonormal Matrices are Cool

- Recall the formula for a projection matrix

$$P = A(A^T A)^{-1} A^T$$

- If we have Q (orthonormal columns) and we want to project Q onto its column space

$$P = Q(Q^T Q)^{-1} Q^T$$

since, by definition $Q^T Q = I$, this reduces to

$$P = QQ^T$$

- Recall the formula for
$$\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$

- If we have Q (orthonormal columns) and we want to solve

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

since, by definition $Q^T Q = I$, this reduces to

$$\hat{\boldsymbol{x}} = Q^T \boldsymbol{b}$$

this also gives us
$$\hat{x}_i = \boldsymbol{q}_i^T \boldsymbol{b}$$

Gram-Schmidt Process

- A process to create orthonormal vectors from independent vectors.

- The idea is to start with one vector and take it as given. Then add another vector and subtract from every new vector its projections in the direction already set. Continue this process for every vector that is added. The result is a set of n orthogonal vectors. Next, divide each vector by its length to end with n orthonormal vectors, $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$.

- The matrix that connects A to Q is the upper triangular matrix R. That is, $A = QR$.

Gram-Schmidt Process with 3 Vectors

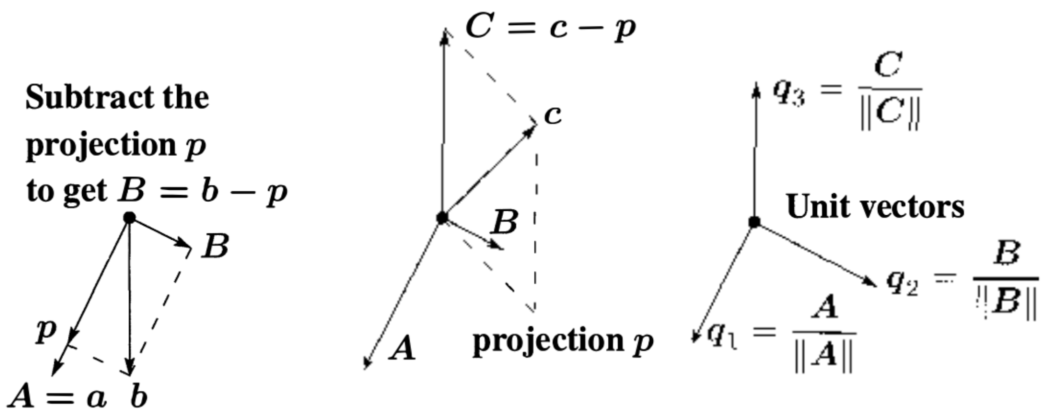

Figure 5: First project $\boldsymbol{b}$ onto the line through $\boldsymbol{a}$ and find the orthogonal $\boldsymbol{B}$ as $\boldsymbol{b} - \boldsymbol{p}$. Then project $\boldsymbol{c}$ onto the $\boldsymbol{AB}$ plane and find $\boldsymbol{C}$ as $\boldsymbol{c} - \boldsymbol{p}$. Divide by $\|\boldsymbol{A}\|, \|\boldsymbol{B}\|, \|\boldsymbol{C}\|$.

- Given 3 independent vectors, $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, we aim to construct 3 orthogonal vectors $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, and then divide by their lengths to produce orthonormal vectors

$$\boldsymbol{q}_1 = \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|}$$
$$\boldsymbol{q}_2 = \frac{\boldsymbol{B}}{\|\boldsymbol{B}\|}$$
$$\boldsymbol{q}_3 = \frac{\boldsymbol{C}}{\|\boldsymbol{C}\|}$$

- Begin by letting $\boldsymbol{A} = \boldsymbol{a}$ (this first direction is accepted, and we will make the other vectors perpendicular to it).

- Next, we must make $\boldsymbol{B}$ such that it is perpendicular to $\boldsymbol{A}$. To do so, start with $\boldsymbol{b}$ and subtract its projection along $\boldsymbol{A}$. That is,

$$\boldsymbol{B} = \boldsymbol{b} - P\boldsymbol{A} = \boldsymbol{b} - \frac{\boldsymbol{A}^T\boldsymbol{b}}{\boldsymbol{A}^T\boldsymbol{A}}\boldsymbol{A}$$

- Next, we must make $\boldsymbol{C}$ such that it is perpendicular to both $\boldsymbol{A}$ and $\boldsymbol{B}$. To do so, start with $\boldsymbol{c}$ and subtract its projection along $\boldsymbol{A}$ and along $\boldsymbol{B}$. That is,

$$\boldsymbol{C} = \boldsymbol{c} - P_A\boldsymbol{A} - P_B\boldsymbol{B} = \boldsymbol{c} - \frac{\boldsymbol{A}^T\boldsymbol{c}}{\boldsymbol{A}^T\boldsymbol{A}}\boldsymbol{A} - \frac{\boldsymbol{B}^T\boldsymbol{c}}{\boldsymbol{B}^T\boldsymbol{B}}\boldsymbol{B}$$

- In this example

$$A = QR$$

$$\begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_3 \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1^T\boldsymbol{a} & \boldsymbol{q}_1^T\boldsymbol{b} & \boldsymbol{q}_1^T\boldsymbol{c} \\ & \boldsymbol{q}_2^T\boldsymbol{b} & \boldsymbol{q}_2^T\boldsymbol{c} \\ & & \boldsymbol{q}_3^T\boldsymbol{c} \end{bmatrix}$$

Gram-Schmidt Process Example with 3 Vectors

- Suppose the independent, non-orthogonal vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \ \boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}, \ \boldsymbol{c} = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$$

- Then $\boldsymbol{A} = \boldsymbol{a}$ has $\boldsymbol{A}^T\boldsymbol{A} = 2$ and $\boldsymbol{A}^T\boldsymbol{b} = 2$. Subtract from $\boldsymbol{b}$ its projection $\boldsymbol{p}$ along $\boldsymbol{A}$

$$\boldsymbol{B} = \boldsymbol{b} - \frac{\boldsymbol{A}^T\boldsymbol{b}}{\boldsymbol{A}^T\boldsymbol{A}}\boldsymbol{A} = \boldsymbol{b} - \frac{2}{2}\boldsymbol{A} = \boldsymbol{b} - \boldsymbol{A} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

  Check that $\boldsymbol{A}^T\boldsymbol{B} = 0$ as required.

- Now, subtract the projections of $\boldsymbol{c}$ on $\boldsymbol{A}$ and $\boldsymbol{B}$ to get $\boldsymbol{C}$

$$\boldsymbol{C} = \boldsymbol{c} - \frac{\boldsymbol{A}^T\boldsymbol{c}}{\boldsymbol{A}^T\boldsymbol{A}}\boldsymbol{A} - \frac{\boldsymbol{B}^T\boldsymbol{c}}{\boldsymbol{B}^T\boldsymbol{B}}\boldsymbol{B} = \boldsymbol{c} - \frac{6}{2}\boldsymbol{A} + \frac{6}{6}\boldsymbol{B} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

  Check that $\boldsymbol{C}$ is perpendicular to both $\boldsymbol{B}$ and $\boldsymbol{C}$ as required.

- Finally, divide $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ by their lengths, which are

$$\|\boldsymbol{A}\| = \sqrt{2}, \ \|\boldsymbol{B}\| = \sqrt{6}, \ \|\boldsymbol{C}\| = \sqrt{3}$$

  Doing so we get

$$\boldsymbol{q}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \ \boldsymbol{q}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \ \boldsymbol{q}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

5 Determinants

5.1 3 by 3 Determinants and Cofactors

Determinant Basics

- Needed for eigenvalues.
- A number associated with every square matrix only.
- Commonly written $\det A = |A|$
- A matrix is invertible when the determinant is nonzero.
- A matrix is singular when the determinant is zero.

Cofactor Formula

- For an $n \times n$ matrix A, the cofactor uses $n!$ terms, have of which are positive, half of which are negative.

- The cofactor formula is a way of connecting an $n \times n$ determinant to an $(n-1) \times (n-1)$ determinant.
- Each summed matrix will have one entry from each row and each column.
- For the i, j cofactor C_{ij}, start by removing row i and column j from the matrix A. C_{ij} equals $(-1)^{i+j}$ time the determinant of the remaining matrix (size $n-1$).
- The cofactor formula along row i is $\det A = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$.
- Useful when the martix entries are mostly 0, then there are few cofactors to find.

5.2 Computing and Using Determinants

Determinant Properties

1. $\det I = 1$
2. Row exchange reverse the sign of the determinant.
   - The determinant of a permutation matrix P is either 1 or -1.
   - From the identity matrix, an odd amount of row exchanges results in a determinant of -1, while an even amount of row exchanges results in a determinant of 1.
3. (a) If we mutiply 1 row by t and leave the other $n-1$ rows alone, the factor t comes out.
   $$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$
   (b) The determinant is linear for each row. It behaves like a linear function on the rows of the matrix.
   $$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$$
4. If 2 rows are equal, the determinant of the matrix is 0.

   This is because of property 2, the exchange rule. On the one hand, exchanging the two identical rows does not change the determinant. On the other hand, exchanging the two rows changes the sign of the determinant. Therefore the determinant must be 0.
5. If $i \neq j$, subtracting t times row i from row j doesn't change the determinant.

   In 2 dimensions, this argument would look like
   $$\begin{vmatrix} a & b \\ c-ta & d-tb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b \\ -ta & -tb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - t\begin{vmatrix} a & b \\ a & b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - t(0) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$
6. If A has a row that is all zeros, then its determinant is 0. We get this property from property 3(a) by letting $t = 0$.
7. The determinant of an upper triangular matrix is the product of the diagonal entries (pivots) $d_1 d_2 \ldots d_n$.

   Property 5 tells us that the determinant of the triangular matrix won't change if we use elimination to convert it to a diagonal matrix with the entries d_i on its diagonal. Then property 3 (a) tells us that the determinant of this diagonal matrix is the product $d_1 d_2 \dots d_n$ times the determinant of the identity matrix (we can factor out each d_i 1 by 1). Property 1 completes the argument. Note that we cannot use elimination to get a diagonal matrix if one of the d_i is zero. In that case elimination will give us a row of zeros and property 6 gives us the conclusion we want.

8. The determinant of matrix A is 0 when A is singular.

   If A is singular, then we can use elimination to get a row of zeros, and property 6 tells us that the determinant is zero. If A is not singular, then elimination produces a full set of pivots $d_1, d_2, \dots, d_n$ and the determinant is $d_1 \dots d_n \neq 0$ (with minus signs from row exchanges).

   We now have a very practical formula for the determinant of a non-singular matrix. In fact, the way computers find the determinants of large matrices is to first perform elimination (keeping track of whether the number of row exchanges is odd or even) and then multiply the pivots.

9. det $AB = (\det A)(\det B)$

   This is very useful. Although the determinant of a sum does not equal the sum of the determinants, it is true that the determinant of a product equals the product of the determinants.

   For example
   $$\det A^{-1} = \frac{1}{\det A}$$
   since
   $$A^{-1}A = I$$
   $$(\det A^{-1})(\det A) = 1$$
   Also,
   $$\det A^2 = (\det A)^2$$
   $$\det 2A = 2^n \det A$$

10. $\det A^T = \det A$

2 by 2 Determinants

- The determinant of the 2 by 2 matrix A is
  $$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

5.3 Areas and Volumes by Determinants

Geometric Interpretation of the Determinant

- The absolute value of a matrix is equal to the volume of the parallelepiped spanned by the row vectors of that matrix.