1 Linear Systems

1.1 Linear Equations

A linear equation with n variables in standard form is:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

It is homogeneous if b = 0

1.2 Linear Systems

A system of linear equations/linear system consists of n variables and m equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{cases}$$

1.3 General Solution

The general solution of a linear system captures all possible solutions of the linear system.
When there are inf. many solutions, we can assign variables as parameters.

Example general solution

$$\left(\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 0 & -2 & 0 & 3 \\ 0 & 1 & 4 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

Assigning the arbitrary variable t to x_3, we have the general solution:

$$\begin{aligned} x_1 &= 3 + 2t \\ x_2 &= -1 - 4t \\ x_3 &= t \\ x_4 &= 2 \end{aligned}$$

Note that you can work backwards from a general solution to get an augmented matrix in RREF.

1.4 Inconsistency/No Solution

A linear system that is *inconsistent* has no solution. Otherwise, if it has at least one solution, it is consistent.

1.5 Augmented Matrices

A linear system can be represented as an augmented matrix. Its dimensions are m x n, where m is the number of rows and n is the number of columns.

$$\left(\begin{array}{cccc|c} x_1 & x_2 & \cdots & x_n & \\ \hline a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

(Top row was added to show how each column in the matrix represents a variable in the system)

Matrix indexing
The (a,b) entry of a matrix is the item in the a^{th} row and b^{th} column.

1.5.1 Solving matrix equations

If we want to find Ax = b, we are essentially finding a coefficient vector x such that the transformation x on A results in the vector b.
If A is invertible, we can also take: x = $A^{-1}b = x$

1.6 Row-echelon form

An augmented matrix is in row-echelon form(REF) if
1. If zero rows exist, they are at the bottom of the matrix
2. The leading entries are further to the right as we move down the rows.

Example:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 2 & 9 \\ 0 & 1 & 4 & 2 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{array}\right)$$

$$\uparrow$$
non-pivot

A column containing a leading entry is called a pivot column. Otherwise, it is a non-pivot column.

1.7 Reduced Row-Echelon form

The augmented matrix is in RREF if
3. The leading entries are 1.
4. In each pivot column, all entries except the leading entry is 0.

Example:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

1.8 Elementary Row Operations

There are 3 types of elementary row operations:

1. Exchanging two rows, $R_i \leftrightarrow R_j$
2. Adding a scalar multiple of a row to another $R_i + cR_j, c \in \mathbb{R}$
3. Multiplying the row by a nonzero constant

Note that performing elementary row operations preserves the solutions of the matrix.

Row Equivalence
Two elementary matrices have the same solutions if their augmented matrices are row equivalent, that is, one can be obtained from the other by performing a set of elementary row operations

Reversing Row Operations
Elementary row operations can be reversed.

1. The reverse of exchanging two rows $R_i \leftrightarrow R_j$ is itself.
2. The reverse of adding a multiple of a row to another, $R_i + cR_j$, is subtracting the multiple of that row: $R_i - cR_j$.
3. The reverse of multiplying a row by a nonzero constant, aR_j, is multiplying the reciprocal: $\frac{1}{a}R_j$

1.9 Row Reduction

An algorithm to reduce rows is described below

Gaussian Elimination
1. Locate the leftmost column that does not consist entirely of zeroes.
2. Interchange the top row with another row to bring a nonzero entry to the top of the column.
3. For each row below the top row, subtract a suitable multiple of the top row so that the entry below becomes zero.
4. Now, cover the top row and start with step 1 again.

Step 1 & 2: Locate non-zero then swap

$$\left(\begin{array}{ccc|c} 0 & 2 & 3 & 8 \\ 4 & 2 & 2 & 7 \\ 2 & 7 & 3 & 1 \end{array}\right) \xrightarrow{R_2 \leftrightarrow R_1} \left(\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 0 & 2 & 3 & 6 \\ 2 & 7 & 3 & 1 \end{array}\right)$$

Step 3: Subtract multiple to make bottom entries 0

$$\left(\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 0 & 2 & 3 & 6 \\ 2 & 7 & 3 & 1 \end{array}\right) \xrightarrow{R_3 - \frac{1}{2}R_1} \left(\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 0 & 2 & 3 & 6 \\ 0 & 6 & 1 & -3 \end{array}\right)$$

Step 4: Cover the top row and repeat steps 1-3

$$\left(\begin{array}{cc|c} 2 & 3 & 6 \\ 6 & 1 & -3 \end{array}\right)$$

Gaussian-Jordan Elimination

Jordan adds two key steps to get the full RREF. Not sure why Gauss didn't think of it.

5. Multiply a suitable constant to each row such that all leading entries become 1

6. Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to make them zero

Step 5: Make the leading entries 1

$$\left(\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 0 & 2 & 3 & 6 \\ 0 & 0 & -8 & -3 \end{array}\right) \xrightarrow{\frac{1}{4}R_1, \frac{1}{2}R_2, -\frac{1}{8}R_3} \left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{2} & 4 \\ 0 & 1 & \frac{3}{2} & 3 \\ 0 & 0 & 1 & \frac{3}{8} \end{array}\right)$$

Step 6: Starting from the bottom, make the upper entries 0.

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{2} & 4 \\ 0 & 1 & \frac{3}{2} & 3 \\ 0 & 0 & 1 & \frac{3}{8} \end{array}\right) \xrightarrow{R_2 - \frac{3}{2}R_3,\ R_1 - \frac{1}{2}R_3} \left(\begin{array}{ccc|c} 1 & \frac{1}{2} & 0 & \frac{61}{16} \\ 0 & 1 & 0 & \frac{39}{16} \\ 0 & 0 & 1 & \frac{3}{8} \end{array}\right)$$

(I should have probably used nicer numbers, but the world isn't nice.)

MATLAB tip

Get the RREF of matrices through MATLAB using the following:

```
% Set formatting to rational for
% fractions if you want
format("Rational")
% define variables if you need to
syms a,b,c
% create the augmented matrix
A = [0 2 3 8; 4 2 2 7; a 7 3 1;]
% call rref function
rref(A)
```

A note on Gaussian-Jordan eliminations with variables
If the matrix you're working with has unknown variables (such as a), try to firstly avoid multiplying the reciprocal, e.g. $\frac{1}{a}R_i$. If you do do this, ensure to account for when $a = 0$, where $\frac{1}{a}$ is not well-defined.

2 Matrix Algebra

2.1 Types of Matrices

2.1.1 Vectors

An n x 1 matrix is called a column vector, and a 1 x n matrix is called a row vector

Column Vector	Row Vector
$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$	$\begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}$

2.1.2 Zero Matrices

Denoted $0_{m\times n}$, it is an m x n matrix with all entries equal to 0.

3 x 4 zero matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

2.1.3 Diagonal Matrix

A Matrix where all non-diagonals are 0.

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$$

d_x can be any number, including 0.

2.1.4 Scalar Matrix

Similar to diagonal matrix, but all elements along the diagonal are the same.

$$\begin{pmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{pmatrix}$$

c must be the same number across the diagonal.

2.1.5 Identity Matrix

All elements across the diagonal must be 1

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2.1.6 Upper Triangular Matrix

Informally, a matrix that forms a triangle along the diagonal, upwards.
$a_{ij} = 0$ for all i > j

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

2.1.7 Strictly Upper Triangular Matrices

Informally, a matrix that forms a triangle *above* the diagonal.
$a_{ij} = 0$ for all i $\geq$ j

$$\begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{pmatrix}$$

2.1.8 Lower Triangular Matrix

Informally, a matrix that forms a triangle along the diagonal, downwards.
$a_{ij} = 0$ for all i < j

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

2.1.9 Strictly Lower Triangular Matrix

Informally, a matrix that forms a triangle *below* the diagonal.
$a_{ij} = 0$ for all i $\leq$ j

$$\begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}$$

2.1.10 Notes on triangular matrices

For any triangular matrix of the same type, strict or not, their sums and products are of the same type.

2.1.11 Symmetric Matrices

For any i, j, $A_{ij} = A_{ji}$. The diagonal can be any number.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

Note that if A is symmetric,then A^2 is symmetric

2.2 Matrix Equality

Two matrices are equal iff they are the same size and their corresponding entries are equal.

2.3 Scalar Matrix Multiplication

A scalar constant λ can be multiplied to a matrix.

$$\lambda \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{pmatrix}$$

Where λ is any real number.

2.4 Matrix Addition

Matrices can be added to each other, but is only defined between matrices of the same size.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \end{pmatrix}$$

2.5 Properties of Matrix Addition and Scalar Multiplation

1. Commutative: A + B = B + A
2. Associative: A + (B + C) = (A + B) + C
3. Additive Identity: $0_{m\times n}$ + A = A
4. Additive Inverse: A + (-A) = $0_{m\times n}$
5. Distributive Inverse a(A+B) = $aA + aB$
6. Scalar addition a+b(A) = $aA + bA$
7. Associative (ab)A = $a(bA)$
8. if aA = $0_{m\times n}$, then either a = 0 or A = 0

2.6 Matrix Multiplication

Symbolically, AB = $(a_{ij})_{m\times p}(b_{ij})_{p\times n} = (\sum_{k=1}^{p} a_{ik}b_{kj})_{m\times n}$. But what the hell does that mean?

1. If $A_{m\times b}$, and $B_{p\times n}$, then the resultant matrix is $AB_{m\times n}$
2. The $(i, j)^{th}$ item of AB is equal to:

   The summation of $A_{ik} * B_{kj}$ for all k in the enumeration of p (the common factor in the size of the matrix.)

2.6.1 An example of matrix multiplication

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Then, the $(1, 2)^{th}$ item of AB is:

$$a_{11} * b_{12} + a_{12} * b_{22} + a_{13} * b_{32}$$

2.6.2 Visualizing matrix multiplication

Localise the row and column vector. Looking at the example above, the $(1, 2)^{th}$ item,

Take the first row of the A matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}$$

Take the second column of the B matrix

$$\begin{pmatrix} b_{12} \\ b_{22} \\ b_{32} \end{pmatrix}$$

Rotate B anti-clockwise, overlay, and add.

$$\begin{matrix} (a_{11} & a_{12} & a_{13}) \\ * & * & * \\ (b_{12} & b_{22} & b_{32}) \\ \| & \| & \| \\ a_{11}b_{12} + & a_{12}b_{22} & + a_{13}b_{32} \end{matrix}$$

Repeat for all other i, j. Suffer.

2.6.3 Properties of Matrix Multiplication

1. Matrix multiplication is non-commutative. That is, AB not necessary equals BA.
   (a) If multiplying A to the left of B, AB, it is called pre-multiplying A to B.
   (b) If multiplying A to the right of B, BA, it is called post-multiplying A to B.
2. Associativity (AB)C = A(BC)
3. Left Distributive Law A(B+C) = AB + AC
4. Right Distributive Law (A+B)C = AC + BC
5. Commutes with scalar multiplication cAB = (cA)B = A(cB)
6. Multiplicative identity For any m $\times$ n matrix A, I_mA = A = AI_n
7. Nonzero Zero Divisor There exists A$\neq$ $0_{m\times p}$ and B$\neq$ $0_{p\times n}$ such that AB = $0_{m\times n}$

8. Zero Matrix For any m×n matrix A, $A0_{n\times p} = 0_{m\times p}$ and $0_{p\times m}A = 0_{p\times n}$

9. Powers of Matrices (Only defined for square matrices!)

   (a) $A^0 = \text{I}$

   (b) $A^n = AA^{n-1}$ *(Essentially, recursively multiply)*

2.7 Transposition of Matrices

The transpose of a $m \times n$ Matrix A, A^T is the $n \times m$ matrix whose (i, j) entry is the (j, i) entry of A.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \to \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

(It doesn't visualize as nice in a non-square matrix but the idea holds)

MATLAB tip

Get the transpose of matrices through MATLAB using the following:

```
% define variables if you need to
syms a,b,c
% create the augmented matrix
A = [0 2 3 8; 4 2 2 7; a 7 3 1;]
% transpose
transpose(A)
```

2.7.1 Properties of Transposition

1. Transpose of a transpose $(A^T)^T = \text{A}$

2. Transpose of a scalar $(cA)^T = cA^T$

3. Distributivity $(A + B)^T = A^T + B^T$

4. Distributivity under multiplication $(AB)^T = B^T A^T$

2.8 Matrix & Vector Equations

We've seen how linear systems can be expressed as augmented matrices. With the properties of multiplication, we can also express them as a matrix equation:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad Ax = b$$

or a vector equation:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b$$

In this notation, a_i (the column vector) is called the coefficient vector for the variable x_i

Example

The linear system:

$$\begin{cases} 3x + 2y - z = 1 \\ x + 2y + z = 3 \\ x + z = 2 \end{cases}$$

can be written as a matrix equation:

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

or a vector equation:

$$x \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

2.9 Homogeneous Linear Systems

A homogeneous linear system, $\mathrm{Ax} = 0$ is always consistent

The zero solution, where $\mathrm{x} = 0_{p \times n}$ is called the trivial solution

The non-zero solution is called the nontrivial solution. *If* the homogeneous linear system has the nontrivial solution, it has infinitely many solutions.

If the trivial solution is a solution to the linear system, then it must be a homogeneous linear system.

$$a_1 x_1 + a_2 x_2 + ... + a_n x_n = b$$
$$0x_1 + 0x_2 + ... + 0x_n = 0$$

In a homogeneous system, if there is **only** the trivial solution, then the columns are linearly independent.

2.9.1 Superposition Principle

If we have the general solution of a system Ax=0 (v_h), and we have any particular solution to **Ax=b** (x_p), then,

$$\text{General Solution of Ax=b} \rightarrow \text{x} = x_p + v_h$$

2.9.2 Lemmas based on properties of Matrix Multiplication

Solutions of linear and homogeneous system
If v is a particular solution to Ax = b, and u is a particular solution to the homogeneous system Ax=0 (same coeff matrix A), then v + u is also a solution to Ax = b.

Proof
By the distributive property of matrix multiplication,

$$A(v + u) = Av + Au = b + 0 = b$$

Solutions of a system minused by itself
If v_1 and v_2 are solutions to the linear system Ax = b, then $v_1 - v_2$ is a solution to the homogeneous linear system Ax = 0.
Proof
Again, by the left distributive law,

$$A(v_1 - v_2) = Av_1 - Av_2 = b - b = 0$$

2.10 Submatrices

A $p \times q$ submatrix of a $m \times n$ matrix A, is formed by taking a $p \times q$ block of the entries of A.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

(Each color corresponds to a possible 2x2 submatrix)

2.11 Block Multiplication

Using submatrices, we can define multiplication on a specific "block" of a matrix.

Formally, Let A be an $m \times p$ matrix and B a $p \times n$ matrix. Let A_1 be a $(m_2 - m_1 + 1) \times p$ submatrix of A obtained by taking rows m_1 to m_2, and b_1 a $p \times (n_2 - n_1 + 1)$ submatrix of B obtained by taking columns n_1 to n_2. Then the product $A_1 B_1$ is a $(m_2 - m_1 + 1) \times (n_2 - n_1 + 1)$ submatrix of AB obtained by taking rows m_1 to m_2 and columns n_1 to n_2.

But what the hell does that mean?

2.11.1 Visualizing Block Matrix Multiplication

Given a 3×2 matrix A, and a 2×3 matrix B, we know that the resultant matrix of AB is 3×3. How do we get the last two rows and columns (a 2x2 submatrix) of AB? Block matrix multiplication is how.

AB:

$$\begin{pmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{pmatrix}$$

We can take the second and third rows of A, and the second and third columns of B to find what the second and third rows and columns (a 2x2 submatrix) of AB is.

A:

$$\begin{pmatrix} ? & ? \\ 4 & 7 \\ 1 & 9 \end{pmatrix}$$

B:

$$\begin{pmatrix} ? & \boxed{\begin{matrix} 5 & 8 \\ 3 & 6 \end{matrix}} \\ ? & \end{pmatrix}$$

$AB_{2\times 2}$:

$$\begin{pmatrix} 5 & 8 \\ 3 & 6 \end{pmatrix} \times \begin{pmatrix} 4 & 7 \\ 1 & 9 \end{pmatrix} = \begin{pmatrix} 28 & 107 \\ 18 & 75 \end{pmatrix}$$

Thus it can be seen that by localizing the correct submatrices, we can calculate the resultant submatrix.

2.11.2 Using block multiplication to solve matrix equations

If we have a coefficient matrix A, how do we find a 3×3 matrix X s.t. AX = b:

$$\begin{pmatrix} 4 & 7 & 3 \\ 8 & 2 & 3 \\ 1 & 9 & 3 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 5 \\ 1 & 4 & 8 \\ 3 & 7 & 5 \end{pmatrix}$$

Because of block multiplication, we know we can take the 3×3 submatrix of A and a 3×1 submatrix of X to find out what the corresponding column vector (e.g. $\begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$) is.

We do this by arranging it in a augmented matrix, which we can stack to solve at once.

$$\left(\begin{array}{ccc|c|c|c} 4 & 7 & 3 & 3 & 2 & 5 \\ 8 & 2 & 3 & 1 & 4 & 8 \\ 1 & 9 & 3 & 3 & 7 & 5 \end{array}\right) \to \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 3 & 9 & 7 \\ 0 & 0 & 1 & 4 & 4 & 5 \end{array}\right)$$

Finding the RREF gives you the X matrix on the right hand side.

2.11.3 Using block multiplication to find unknowns

This is a natural extension from the previous example, but more complicated. Say we have matrixes A, B, and C s.t.

$$\underset{A}{\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}} * \underset{B}{\begin{pmatrix} 1 & 21 & 3 & 4 \\ 2 & 21 & 5 & 6 \\ 7 & 21 & 8 & 9 \end{pmatrix}} = \underset{C}{\begin{pmatrix} 10 & ? & 12 & 13 \\ 14 & ? & 16 & 17 \\ 18 & ? & 20 & 21 \end{pmatrix}}$$

We can use block multiplcation to create equations to solve for A. (e.g. $x_{11} + 2x_{12} + 7x_{13} = 10$). After getting A, we can multiply it to B to get the unknown column in C.

2.12 Inverse of Matrices

For a matrix to have an inverse, it must be such that $AA^{-1} = I = A^{-1}A$. From this, it can be easily seen that inverse for non-square matrices is **not** well-defined!.

So we define invertibility as such:
A $n \times n$ square matrix is invertible if there exists a matrix **B** such that

$$AB = I_n = BA$$

Invertibility is an important rule on matrices, and given invertibility there are many statements that we can draw on the matrix.

2.12.1 Computing an Inverse

We can easily compute an inverse. Given that $AA^{-1} = I$, We can follow the same ideas as described in 2.11.2 to compute the inverse:

$$(A \mid I) \xrightarrow{\text{RREF}} (I \mid A^{-1})$$

MATLAB tip

Get the inverse of square matrices through MATLAB using the following:

```
% Set formatting to rational for
% fractions if you want
format("Rational")
% define variables if you need to
syms a,b,c
% create the augmented matrix
A = [0 2 3; 4 2 2; a 7 3;]
% call inverse function
inv(A)
```

2.12.2 Laws on Invertible Matrices

All the following laws assume an $n \times n$ square matrix unless otherwise stated.

Law #1: Inverse of 2×2 Square Matrices

A 2x2 square matrix A = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible iff $ad - bc \neq 0$ and its inverse is given:

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Law #2: Cancellation Law of an invertible matrix

1. Left Cancellation: if B and C are $n \times m$ matrices with $AB = AC$, then $B = C$

2. Right Cancellation: if B and C are $m \times n$ matrices with $AB = AC$, then $B = C$

Law #3: Consistency of Invertible Matrix and uniqueness of solution
For any $n \times 1$ matrix b, $Ax = b$ has a unique solution.

Law #4: Homogeneous Invertible System solution
Given the invertible homogeneous linear system $Ax = 0$, we can conclude that the trivial solution is the only solution.

2.12.3 Properties of Inverses & Equivalent Statements of Invertibility

Given that A is invertible, (not every square matrix is invertible!),

1. $(A^{-1})^{-1} = A$
2. For any nonzero real number a $\in \mathbb{R}$, aA is invertible with inverse $aA^{-1} = \frac{1}{a}A^{-1}$
3. A^T is invertible with inverse $(A^{-1})^T$
4. if B is an $n \times n$ invertible matrix, then (AB) is invertible with inverse $\mathbf{B^{-1}A^{-1}}$
    (a) Extending this law, if $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$
5. A is invertible iff A is a product of elementary matrices
6. A is invertible iff the RREF of A is the identity matrix
7. A is invertible iff the homogeneous system Ax=0 has only the trivial solution.
8. A is invertible iff Ax=b has a unique solution for all b.
9. A is invertible iff it has a left inverse, i.e. BA = I_n
10. A is invertible iff it has a right inverse, i.e. AB = I_n
11. A is invertible iff det(A) $\neq 0$
12. A is invertible iff the columns/rows of A spans $\mathbb{R}^n$
13. A is invertible iff the columns/rows of A are linearly independent
14. A is invertible iff A has full rank
15. A is invertible iff nullity(A) = 0

16. A is invertible iff 0 is not and eigenvalue of A.

17. A is invertible iff the linear transformation T defined by A is injective and surjective

2.13 Elementary Matrices

An $n \times n$ matrix n is called an elementary matrix if it can be obtained from the identity matrix I_n by performing a single elementary row operation

Performing row operations on any matrix is equivalent to premultiplying the corresponding elementary matrix. That is:

$$A \xrightarrow{r} EA$$

Where E is the elementary matrix corresponding to operation r.

Applying that repeatedly, we can define row equivalent matrices like so:

$$A \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} B$$

$$B = E_k ... E_2 E_1 A$$

Manipulating Elementary Matrices
If we know the RREF of A is I, that is, $A \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} I$, then,

$$I = E_k ... E_1 A$$

We can make A the subject by premultiplying the inverse of each matrix onto I sequentially:

$$A = E_1^{-1} ... E_k^{-1} I = E_1^{-1} ... E_k^{-1}$$

Taking it further, we can derive A^{-1} by taking the inverse of both sides and applying property #4 and #1 of inverses:

$$A^{-1} = (E_1^{-1} ... E_k^{-1})^{-1} = E_k ... E_2 E_1$$

2.13.1 Inverse of Elementary Matrices

Every elementary matrix is invertible, and corresponds to the reverse of the row operation corresponding to E.

1. $$I_n \xrightarrow{R_i + cR_j} E \quad \xrightarrow{R_i - cR_j} I_n \quad \Rightarrow \quad E : R_i + cR_j, \quad E^{-1} : R_i - cR_j.$$

2. $$I_n \xrightarrow{R_i \leftrightarrow R_j} E \quad \xrightarrow{R_i \leftrightarrow R_j} I_n \quad \Rightarrow \quad E : R_i \leftrightarrow R_j, \quad E^{-1} : R_i \leftrightarrow R_j.$$

3. $$I_n \xrightarrow{cR_i} E \quad \xrightarrow{\frac{1}{c}R_i} I_n \quad \Rightarrow \quad E : cR_i, \quad E^{-1} : \frac{1}{c}R_i.$$

2.13.2 Elementary Matrices and Unit Lower Triangular Matrices

See here for definition on unit triangular matrix.

For a row operation $R_i + cR_j$ where **i** $>$ **j**, it's corresponding elementary matrix is a lower triangular matrix

E^{-1} is also a unit triangular matrix.

Then, if $E_1, E_2, E_3, ..., E_k$ are row operations of the type $R_i + cR_j$ where $i > j$, then the $E_1E_2E_3..E_k$ and $E_1^{-1}E_2^{-1}E_3^{-1}..E_k^{-1}$ is a unit lower triangular matrix as well.

2.13.3 Determinant of Elementary Row Operations

This section is repeated later but it's here for ease

Transformation	Effect on Determinant
$A \xrightarrow{R_i+aR_j} B$	$\det(B) = \det(A)$
$A \xrightarrow{cR_i} B$	$\det(B) = c\det(A)$
$A \xrightarrow{R_i \leftrightarrow R_j} B$	$\det(B) = -\det(A)$

2.14 LU Factorization

LU factorization is the process of describing a matrix as a product of a unit lower triangular matrix and a row-echelon of A. We can write $A = LU$, where L is the lower triangular matrix, and U is a row-echelon form of A.

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix}$$

The LU factorization of A is:

$$A = LU$$

where:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 6 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

Note that not every matrix has a LU factorization. Additionally, A is **row-equivalent** to U.

Notice that LU factorization is just doing gaussian elimination on matrix. If there are situations in which you cannot form a unit lower triangular matrix when doing gaussian elimination (e.g. you have to swap a row), then you cannot LU factorize.

2.14.1 Unit Lower Triangular Matrices

Defined as a lower triangular matrix with 1 as the diagonal entries:

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ ? & 1 & 0 & 0 \\ ? & ? & 1 & 0 \\ ? & ? & ? & 1 \end{pmatrix}$$

Where ? is any real number.

A product of any unit triangular matrix is a unit triangular matrix too.

2.14.2 Algorithm to LU Factorize

Using our knowledge on elementary matrices and unit lower triangular matrices, we compute a LU factorization for A.

Suppose

$$A \xrightarrow{r_1, r_2, \ldots, r_k} U$$

where each row operation r_i is of the form $R_i + cR_j$ for some $\mathbf{i} > \mathbf{j}$ and real number c, and U is a row-echelon form of A. Let E_i be the elementary matrix corresponding to r_i, for $i = 1, 2, \ldots, k$. Then

$$E_k \cdots E_2 E_1 A = U \quad \Rightarrow \quad A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} U = LU$$

where

$$L = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

Then

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{pmatrix} \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}$$

is a LU factorization of A.

To quickly compute the L matrix, for each row operation $r_x = R_i + c_x R_j$, simply put -c_x in the (i,j) entry of L.

MATLAB tip

Get the LU factorization using the following. Note that this is LU factorization *without* pivoting. This function will not work if the diagonal coefficients are equal to 0.

```
% (credit to rayryeng on stackoverflow lol)
function [L, U] = lu_nopivot(A)

n = size(A, 1);
L = eye(n);
for k = 1 : n
    L(k + 1 : n, k) = A(k + 1 : n, k) / A(k, k);
    for l = k + 1 : n
        A(l, :) = A(l, :) - L(l, k) * A(k, :);
    end
end
U = A;

end

% Define a matrix A
A = [2 3 1; 4 7 1;6 9 4]

% Call the function
[L,U] = lu_nopivot(A);
% If you're paranoid, verify L * U = a
L * U
```

2.14.3 Solving Matrix Equations using LU Factorizations

Given a linear system Ax=b, and a LU factorization of A, the linear system can be re-written as $\mathbf{LUx} = \mathbf{b}$.

1. Let Ux be an arbitrary variable y such that $\mathbf{Ly} = \mathbf{b}$.
2. Using archaic block multiplication knowledge detailed here, solve the system for y.
3. Now, we have a $\mathbf{Ux} = \mathbf{y}$. Apply step 2 but solve for x.
4. Sacrifice a goat

2.15 Determinants

The determinant is a value calculated from a matrix that lets us know if a matrix is invertible among other things.

A matrix with detrminant 0 is called singular.

2.15.1 Determinant by Cofactor Expansion

$$\det(A) = \sum_{k=1}^{n} \underline{(-1)^{i+k}} a_{ik} \det(A_{ik})$$

or

$$\det(A) = \sum_{k=1}^{n} \underline{(-1)^{j+k}} a_{kj} \det(A_{kj})$$

The formula essentially says that you can choose any 1 column or row and expand along that to calculate the determinant.

The sign of cofactor
The orange underlined portion of the formula above is the sign of the cofactor, and can be visualised like so:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Any value along a *diagonal* is positive

Example of cofactor expansion
Given a matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

The determinant of A can be computed by cofactor expansion along the first row:

$$\det(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

This involves calculating the 2x2 determinants:

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

2.15.2 Determinant by Reduction

Determinant by reduction is built on the fact that the determinant of a triangular matrix is the product of its diagonal entires.

Knowing this, we can manipulate any matrix to make it a triangular matrix and then multiply the diagonal entries to get the final determinant.

Determinant of Elementary Row Operations

Transformation	Effect on Determinant
$A \xrightarrow{R_i + aR_j} B$	$\det(B) = \det(A)$
$A \xrightarrow{cR_i} B$	$\det(B) = c\det(A)$
$A \xrightarrow{R_i \leftrightarrow R_j} B$	$\det(B) = -\det(A)$

With the above if A and R are square matrices s.t. $A \xrightarrow{r_1} \xrightarrow{r_2} ... \xrightarrow{r_k} R$,

$$R = E_k ... E_2 E_1 A$$

Then,

$$det(R) = det(E_k)...det(E_2)det(E_1)det(A)$$

2.15.3 Properties of Determinant

1. Invariant under transpose: det(A) = det(A^T)
2. Determinant of Product: Provided A and B are the same size, det(AB) = det(A)det(B)
   (a) Determinant of Squares: det(A^2) = det(A)det(A) = $(\text{det(A)})^2$
3. Determinant of Inverse: Provided A is invertible, det(A^{-1}) = $det(A)^{-1}$
4. Determinant of Scalar Multiplication: det(cA) = $c^n det(A)$, where n is the order of the matrix A

2.15.4 Some Determinants

1. Determinant of a $n \times n$ Identity Matrix: 1
2. Determinant of a *negative* Identity Matrix: $(-1)^n$
3. Determinant of a scalar matrix: c^n where c is the scalar
4. Determinant of a upper triangular matrix: Product of its diagonal

2.16 Adjoint

Similar to transposition, the adjoint of A is the $n \times n$ square matrix whose (i,j) entry is the (j,i) cofactor of A.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{cofactor} \begin{pmatrix} cf_{11} & cf_{12} & cf_{13} \\ cf_{21} & cf_{22} & cf_{23} \\ cf_{31} & cf_{32} & cf_{33} \end{pmatrix} \xrightarrow{transpose} \begin{pmatrix} cf_{11} & cf_{21} & cf_{31} \\ cf_{12} & cf_{22} & cf_{32} \\ cf_{13} & cf_{23} & cf_{33} \end{pmatrix}$$

MATLAB tip

Get the adjoint of matrices through MATLAB using the following:

```
% Define a function to get cofactors
function C = cofactor(A)
    n = size(A, 1);  % Get the size of the matrix
    C = zeros(n);    % Initialize the cofactor matrix

    for i = 1:n
        for j = 1:n
            % Calculate the minor
            % by removing the ith row and jth column
            M = A;
            M(i, :) = [];     % Remove the ith row
            M(:, j) = [];     % Remove the jth column

            % Calculate the cofactor
            C(i, j) = ((-1)^(i + j)) * det(M);
        end
    end
end

% create the augmented matrix
A = [0 2 3 8; 4 2 2 7; 7 7 3 1;]
% get cofactor then transpose
transpose(cofactor(A))
```

2.16.1 Adjoint Formula

If A is a square matrix and adj(A) its adjoint. Then,

$$A(adj(A)) = det(A)I$$

Determinant of a square adjoint
As we have the square formula above, we can use that to calculate a determinant of an adjoint given det(A).

$$\text{det(A(adj(A)))} = \text{det(det(A)I)}$$

det(A) is a scalar value, so det(A)*I is some identity matrix of scalar det(A). Recall the determinant of a triangular matrix is the product of it's diagonal. So, rearranging the RHS,

$$\text{det(det(A)I)} = \text{det(A)}^n$$

Using known properties of determinants, we re-arrange the left hand side to:

$$\text{det(a) * det(adj(A))}$$

Finally, we have

$$\text{det(adj(A))} = \frac{\text{det(A)}^n}{det(A)} = \text{det(A)}^{n-1}$$

2.16.2 Adjoint Formula for inverse

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

3 Euclidian Vector Spaces

Recall the definitions of Vectors:

v = $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

Each entry v_i is also known as the i-th coordinate.
The Euclidian n-space, denoted as $\mathbb{R}^n$ is the collection of all n-vectors

$$\mathbb{R}^n = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \mid v_i \in \mathbb{R} \text{ for } i = 1, \ldots, n \right\}.$$

3.1 Vector Algebra

Note that bolded letters refer to vectors, and non-bolds refer to scalars

1. The sum $\mathbf{u} + \mathbf{v}$ is a vector in $\mathbb{R}^n$
2. Commutative: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. Associative: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
4. Zero vector: $0 + \mathbf{u} = \mathbf{u}$
5. Negative Vector: **-u** is a vector s.t. $\mathbf{u}$ + **-u** $= 0$
6. Scalar Multiple: c**u** is a vector in $\mathbb{R}^n$
7. Distributivity of Scalar:
   (a) a($\mathbf{u} + \mathbf{v}$) = a**u** + a**v**
   (b) (a+b)**u** = a**u** + b**u**
8. Associativity of scalar multiplication: (ab)**u** = a(b**u**)
9. Zero result: If a**u** = 0, then either a = 0 or $\mathbf{u} = 0$

3.2 Abstract Vector Spaces

An abstract vector space is a *mathematical structure* that consists of a set of objects that are not necessarily vectors that satisfy axioms of vector space. These objects can be functions, polynomials, matrices or any kind of mathematical object. They just have to **behave like a vector space**. This allows us to define vector spaces for non-vector objects, and answer exam questions so we don't fail.
*Note that a **real** vector space is a specific instance of abstract vector spaces where the scalars are real numbers. It's not the opposite of an abstract vector space as it may seem.*

A set V equipped with addition and scalar multiplication is said to be a vector space over $\mathbb{R}$ if it satisfies the following axioms:

1. Inclusion of Sum: For any vectors **u** and **v** in V, the sum $\mathbf{u} + \mathbf{v} \in$ V
2. Commutativity: For any vectors **u**, **v** in V, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. Associativity: For any vectors **u**, **v**, **w** in V, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
4. Zero Vector: There is a vector **0** in V such that $0 + \mathbf{v} = \mathbf{v}$ for all vectors in V.
5. Negativity: For any vector **u** in V, there exists a vector **-u** in V s.t. $\mathbf{u}$ + **-u** $= 0$

6. Scaling: For any scalar a in $\mathbb{R}$ and vector $\mathbf{u}$ in V, a$\mathbf{u}$ is a vector in V.
7. Distributivity:
   (a) a($\mathbf{u}$ + $\mathbf{v}$) = a$\mathbf{u}$ + a$\mathbf{v}$
   (b) (a+b)$\mathbf{u}$ = a$\mathbf{u}$ + b$\mathbf{u}$
8. Associativity of scalar multiplication: For any scalars a, b $\in \mathbb{R}$ and vector $\mathbf{u}$ in V, a(b$\mathbf{u}$) = (ab)$\mathbf{u}$
9. Identity: For any vector $\mathbf{u}$ in V, 1$\mathbf{u}$ = $\mathbf{u}$

3.3 Multiplying Vectors (Dot Product)

We can multiply two $n \times 1$ vectors $\mathbf{u}$ and $\mathbf{v}$ ($\mathbf{u} \cdot \mathbf{v}$) by taking the transpose ($\mathbf{u}^T$) of the left vector multiplied by the right vector.
Doing this, the result is:

$$u_1v_1 + u_2v_2 + ... + u_nv_n$$

or symbolically,

$$\textstyle\sum_{i=1}^{n} u_iv_i$$

3.4 Distance length of vector

We can get a length of a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ by using the pythagorean theorum.

$$Distance = \sqrt{x^2 + y^2}$$

Similarly, for a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in $\mathbb{R}^3$, we can compute:

$$Distance = \sqrt{(\sqrt{x^2 + y^2})^2 + z^2} = \sqrt{x^2 + y^2 + z^2}$$

3.5 Norm of a vector

Using the pattern observed in finding a distance of a vector in $\mathbb{R}^2$ and $\mathbb{R}^3$, we can find a distance of <u>any</u> vector through the following:

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2u_2^2 + ... + u_n^2}$$

We call this the norm of a vector.

3.6 Properties involving the inner product and the norm

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ and a, b, c be some scalars

1. Symmetric $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. Scalar Multiplication: $\mathrm{c}\mathbf{u}\cdot\mathbf{v} = (\mathrm{c}\mathbf{u})\cdot\mathbf{v} = \mathbf{u}\cdot(\mathrm{c}\mathbf{v})$
3. Distribution: $\mathbf{u} \cdot (\mathrm{a}\mathbf{v} + \mathrm{b}\mathbf{w}) = \mathrm{a}\mathbf{u}\cdot\mathbf{v} + \mathrm{b}\mathbf{u}\cdot\mathbf{w}$
4. Positive Definite: $\mathbf{u}\cdot\mathbf{u}\geq 0$ if and only if $\mathbf{u} = 0$
5. Norm of scalar multiples: $\|c\mathbf{u}\| = \mid c \mid \times \|\mathbf{u}\|$

3.7 Unit Vectors

A vector $\mathbf{u}$ is a unit vector if its norm is 1.

3.7.1 Normalizing

We can normalize a vector by multiplying it by the reciprocal of its norm.

$$\text{Unit vector of } \mathbf{u} = \mathbf{u} \cdot \frac{1}{\|\mathbf{u}\|}$$

3.7.2 Distance between vectors

The distance between two vectors is the norm of their difference:

$$\text{distance between } \mathbf{u} \text{ and } \mathbf{v} = \|\mathbf{u} - \mathbf{v}\|$$

Plugging this into the formula for norms,

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + ... + (z_1 - z_2)^2}$$

where $x_1 - z_1$ are coefficients in $\mathbf{u}$ and $x_2 - z_2$ are coefficients in $\mathbf{v}$

3.8 Angle between vectors

While we're at this, we can also get the angle between two vectors $\mathbf{u}$ and $\mathbf{v}$ like so:

$$\theta = \cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$$

3.9 Linear Combinations

Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ be vectors in $\mathbb{R}^n$. A linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ is

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k,$$

for some $c_1, c_2, \ldots, c_k \in \mathbb{R}$. The scalars $c_1, c_2, \ldots, c_k$ are called coefficients.

3.10 Linear Spans

The span of a set of vectors is the subset of $\mathbb{R}^n$ containing all the linear combinations of the vectors.

Informally put, it's the subspace/set of points that the vectors can "reach"

3.10.1 Valid Spans

A set S = $\{u_1, ..., u_k\}$ is a valid span of the Euclidean n space if and only if $(u_1 u_2 ... u_k | v)$ is consistent for every $v \in \mathbb{R}^n$, that is, it has no zero rows.

3.10.2 Checking if a vector falls in a span

To check if a vector falls in a span, it is simply checking if the vector is a linear combination of the set of vectors.

1. Let S = $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$
2. Form the $n \times k$ matrix A = $(\mathbf{u}_1 \ \mathbf{u}_2 \ ... \ \mathbf{u}_k)$
3. Set up the equation Ax = v, where v is the vector that is being checked
4. If the system is consistent, the vector is in the span of the vectors. The solutions to the system are the coefficients to the linear combination.

3.10.3 Example of checking for inclusion in span

We want to check if $\mathbf{v} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$ is in the span of $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{u}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.
We form an augmented matrix:

$$\left(\begin{array}{cc|c} 1 & 3 & 7 \\ 2 & 1 & 5 \end{array}\right)$$

Perform row operations to find the RREF.

$$\begin{pmatrix} 1 & 0 & \frac{8}{5} \\ 0 & 1 & \frac{9}{5} \end{pmatrix}.$$

From this, we can see that V can be expressed as a linear combination of u_1 and u_2 like so:

$$\mathbf{v} = \frac{8}{5}\mathbf{u}_1 + \frac{9}{5}\mathbf{u}_2.$$

3.10.4 Standard Basis

A set of n vectors $\{u_1, u_2, ...u_n\}$ is called a standard basis if it spans the whole of $\mathbb{R}^n$

A set of vectors will be a standard basis if it's linearly independent. We can check this by reducing the matrix A $= (u_1 u_2 ... u_n)$ and seeing if there are any non-pivot columns. If there are no non-pivot columns, the vectors are linearly independent and span $\mathbb{R}^n$.

3.10.5 Properties of Linear Spans

1. The zero vector is in the span.
2. The span is closed under scalar multiplication. Any vector $\mathbf{u}$ in span(S) and scalar a, the vector a$\mathbf{u}$ is a vector in span(S).
3. The span is closed under addition. For any vectors $\mathbf{u}$ and $\mathbf{v}$ in span(S), their sum $\mathbf{u} + \mathbf{v}$ is a vector in span(S)
4. Condensing (2) and (3), the span is closed under linear combination.

3.10.6 Subsets of Spans (Span relations)

To check if span$(S) =$ span(T), we check that:

- span$(S) \subseteq$ span(T), that is,

$$\left(\begin{array}{c|c} T & S \end{array}\right) = \left(\begin{array}{cccc|c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_m & \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{array}\right)$$

  is consistent.

- span$(T) \subseteq$ span(S), that is,

$$\left(\begin{array}{c|c} S & T \end{array}\right) = \left(\begin{array}{cccc|c|c|c|c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k & \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_m \end{array}\right)$$

  is consistent.

Example

Let $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \right\}$ and $T = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$.

To check if span$(S) \subseteq$ span(T),

$$\left(\begin{array}{ccc|c|c|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & -1 & 1 & 2 & 3 & 3 \\ 1 & 1 & 1 & 1 & 1 & 2 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right)$$

This is consistent

To check if span$(T) \subseteq$ span(S),

$$\left(\begin{array}{ccc|c|c|c} 1 & 1 & 2 & 1 & 1 & 1 \\ 1 & 2 & 3 & 1 & -1 & 2 \\ 1 & 1 & 2 & 1 & 0 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \color{red}{1} & 0 \end{array}\right)$$

This is <u>not consistent</u>

This shows that $\mathrm{span}(T) \not\subseteq \mathrm{span}(S)$. In particular, $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \notin \mathrm{span}(S)$.

Span(S) $\subseteq$ span(T)

3.11 Solution sets of a linear system

The <u>set</u> of solutions to a linear system Ax = b is a subset in $\mathbb{R}^n$. We can express the set in two ways:

<u>Implicit</u> Expression of set of solutions

$$\mathrm{V} = \{\ u \in \mathbb{R}^n \mid \mathbf{Au{=}b}\ \}$$

writing it out a little more verbosely,

$$\mathrm{V} = \left\{ \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \middle| \; A \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = b \right\}$$

<u>Explicit</u> Expression of set of solutions

$$V = \{\underline{\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_k\mathbf{v}_k} \mid s_1, s_2, \ldots, s_k \in \mathbb{R}\}$$

Where the <u>blue section</u> is the general solution

3.11.1 Example: Writing out solution sets

Considering this linear system:

$$\begin{aligned} 3x + 2y - z &= 0 \\ y - z &= 0 \end{aligned}$$

We write the solution set <u>implicitly</u> as:

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \; 3x + 2y - z = 1,\ y - z = 0 \right\}$$

To write it out explicitly, we must first get the general solution. Do this intuitively, or use an augmented matrix:

$$\text{Augmented matrix:} \quad \begin{pmatrix} 3 & 2 & -1 & | & 1 \\ 0 & 1 & -1 & | & 0 \end{pmatrix}$$

$$\text{RREF:} \quad \begin{pmatrix} 1 & 0 & \frac{1}{3} & | & \frac{1}{3} \\ 0 & 1 & -1 & | & 0 \end{pmatrix}$$

NOTE: that when you are writing the solution space out explicitly, you are finding the nullspace of the RREF.
From this, we can see that the general solution is:
$x = \frac{1}{3}(1-s),\ y = z = s,\ s \in \mathbb{R}$
From the general solution, we write the solution set explicitly as:

$$\left\{ \begin{pmatrix} \frac{1}{3} \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 1 \end{pmatrix} \,\middle|\, s \in \mathbb{R} \right\}$$

3.11.2 Spans of Solution Sets

Writing out a solution set explicitly, we can easily see that a solution set spans a certain space itself. From the example above, the solution set is the span:

$$\text{Span}\left\{ \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 1 \end{pmatrix} \right\} + \begin{pmatrix} \frac{1}{3} \\ 0 \\ 0 \end{pmatrix}$$

This is because of the free parameter s, which is any real number. So the amount of dimensions that a solution set spans is closely linked to the amount of non-pivot columns in the solution. (*this is an intuition that will help later*)

We can also call this a subspace if it fulfils certain criteria.

3.12 Subspaces

Subset V of the Euclidean n-space is a vector space if it satisfies these 3 axioms:

1. V contains the zero vector
2. V is closed under scalar multiplication. For any vector **u** in V and scalar a, the vector a**u** is in V.
3. V is closed under addition. For any **u**, **v** in V, the sum **u**+**v** is in V.

3.12.1 Homogeneous Linear Systems and Subspaces

It turns out that for a linear system **Ax=b**, for its solution set V= { **u** | Au = b} to be a subspace, the system must be homogeneous. That is, Au = 0.
The solution set to a homogeneous system is called a solution space.

3.12.2 Subspaces under union

For 2 subspaces V and W of arbitrary Euclidian space $\mathbb{R}^n$,
$V \cup W$ is a subspace **only if**

$$V \subseteq W \text{ or } W \subseteq V$$

This is because if they are not subsets of each other, then there exists some vector u$\in V$ (resp. W) such that v$\notin W$ (resp. V). This means that closure under addition is not guaranteed

Geometrically, unioning subspaces represents the vectors that belong to V and W, but not their linear combinations.

Basis for combined subspace If B and C are bases for V and W, then $B \cup C$ is a basis for $V + W$, after removing dependent columns. $B + C$ is not the basis, because adding the columns together might introduce dependency.

3.12.3 Subspaces under addition

Two subspaces V and W are **always** subspaces under addition.

1. V+W contains the zero vector; 0+0 = 0
2. V+W is closed under scalar multiplication. (v + w) * 3 = (3v + 3w). Since V and W are closed under scalar multiplication itself, this holds
3. V+W is closed under scalar addition. (3v + 3w) + (21v + 32w) = (24v + 35w). Same as above.

Geometrically, adding subspace represents a subspace that spans all the vectors and linear combinations in V and W.

3.12.4 Subspaces under Intersection

Two subspaces V and W are **always** subspaces under Intersection

1. $V \cap W$ contain the zero vector.
2. $V \cap W$ is closed under addition: $x \in V \cap W \to (x \in V) \cap (x \in W)$. Thus, addition is defined.
3. $V \cap W$ is closed under scalar multiplication.

Geometrically, intersecting two subspaces are the vectors contained in both bases.

Finding subspace of intersections You are essentially finding a subspace s.t. if V = $\{v_1, v_2, v_3\}$ and W = $\{w_1, w_2, w_3\}$

$$x \in V \cap W \to c_1v_1 + c_2v_2 + c_3v_3 = k_1w_2 + k_2w_2 + k_3w_3 = x$$

Simply create an augmented matrix (V|W), then reduce

3.12.5 Describing subsets using linear equations

If you have a matrix U, and you want to describe U using some linear equation $ax_1 + bx_2 + ... + cx_n$, take transpose of U, RREF, and the general solution will describe a linear equation of that form.

3.13 Linear Independence

A set of vectors $\{u_1, u_2, ..., u_k\}$ are linearly independent if and only if the only constants fulfilling the equation

$$c_1u_1 + c_2u_2 + ...c_ku_k = 0$$

are 0. That is to say, each vector contributes a unique value in a dimension (informally). If there exists some constant such that the equation is true, the set is linearly dependent Note: Any set containing the 0 vector is linearly dependent1

3.13.1 Checking for linear independence

By arranging it into a matrix, we are checking if there exists $c_1 - c_n$ such that:

$$\begin{pmatrix} v_{11} & v_{21} & \cdots & v_{k1} \\ v_{12} & v_{22} & \cdots & v_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{kn} \end{pmatrix} * \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

We put it into an augmented matrix and solve for c

$$\begin{pmatrix} v_{11} & v_{21} & \cdots & v_{k1} & | & 0 \\ v_{12} & v_{22} & \cdots & v_{k2} & | & 0 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{kn} & | & 0 \end{pmatrix}$$

After solving, for it to be linearly independent, the system should have no non-pivot columns. That is to say, Ax = 0 only has the trivial solution. If it is linearly dependent, there will be non-pivot columns, suggesting that there are infinitely many values such that the $\sum c_iv_i = 0$

3.13.2 Linear Independence and Adding/Remove vectors

Here are some fairly obvious theorums on linear independence:
1. If $u_1, u_2, .., u_k$ are linearly dependent, adding any vector v to the set will result in a linearly dependent set
2. If $u_1, u_2, .., u_k$ are linearly independent, adding a vector v that is not a linear combination of any vectors $u_1 - u_k$ will result in a linearly independent set
3. If $u_1, u_2, ..., u_k$ are linearly independent, any subset of the set is linearly independent
When manipulating sets for linear independence, you can check if a vector is a linear combination of the vectors of the set by putting them in an augmented matrix.

3.14 Basis

A basis of a space is essentially an unambiguous (all points can only be represented in one way) coordinate system for the space.

A set S = $\{u_1, ..., u_k\} \subseteq V$ is a basis for V iff

1. Span(S) = V
2. S is linearly independent

This says every vector u $\in$ V can be written as a combination of vectors in S. Note that a basis for a subspace is not unique.

3.14.1 Basis of Homogeneous systems

For a homogeneous linear system of n equations, its solution space is $\mathbb{R}^n$

3.14.2 Properties of Bases

Let V be a subspace, and B a basis of that subspace. | B| = k

1. Uniqueness of size: Suppose $S = \{u_1, u_2, ..., u_k\}$ and $T = \{v_1, v_2, ..., v_m\}$ are bases for a subspace. Then, k = m.
2. Size requirements?
   (a) if S = $v_1, v_2, ...v_m$ is a subset of V with m > k, then S is linearly dependent
   (b) if S = $v_1, v_2, ...v_m$ is a subset of V with m < k, then S cannot span V.

3.14.3 Checking if a matrix is a basis

Let V be a k-dimensional subspace of $\mathbb{R}^n$, and dim(V) = k. If dim(V) is not provided, calculate by checking the solution set

Method 1:
Check that the matrix S is linearly independent subset of the subspace V. (Do this by plugging in values and seeing if they are consistent in the given system)
If S has k vectors, | S |=k, then S is a basis for V.

Method 2:
Check that $V \subseteq span(S)$.
If S has k vectors, | S |=k, then S is a basis for V.

3.14.4 Example: Checking if a matrix is a basis

Method 1 example:

$$Let\ V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mid x_1 - 2x_2 + x_3 = 0,\ x_2 + x_3 - 2x_4 = 0 \right\}$$

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\}$$

S is clearly linearly independent. To check if it's a subset of V, we check that the vectors in S fulfill the requirements of V:
Vector 1:

$$\begin{aligned} x_1 - 2x_2 + x_3 = 0 \quad &\rightarrow \quad (1) - 2(1) + (1) = 0 \\ x_2 + x_3 - 2x_4 = 0 \quad &\rightarrow \quad (1) + (1) - 2(1) = 0 \end{aligned}$$

Vector 2:

$$\begin{aligned} x_1 - 2x_2 + x_3 = 0 \quad &\rightarrow \quad (3) - 2(1) + (-1) = 0 \\ x_2 + x_3 - 2x_4 = 0 \quad &\rightarrow \quad (1) + (-1) - 2(0) = 0 \end{aligned}$$

Now, we check that dim(V) = number of vectors in S = 2
Checking dim(V): Express V explicitly

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & 1 & -2 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 3 & -4 \\ 0 & 1 & 1 & -2 \end{pmatrix}$$

We can see that there are 2 non pivot columns, and hence the dimension of V is 2.
Therefore, S is a basis for V.

Method 2 example

$$Let\ T = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \right\}$$

V = span(T)

$$Let\ S = \left\{ \begin{pmatrix} 0 & 0 & 4 & 0 \\ 2 & 1 & 6 & 2 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 4 & 1 \\ -1 & -1 & -3 & -1 \end{pmatrix} \right\}$$

To show that S is a basis for V, we first check that span(T) = V $\subseteq$ span(S). Always check that the "bigger" one, the basis you're checking for, is a subset of the smaller one.We use the method outlined in 3.10.5, creating an augemnted matrix (S|T)

$$\begin{pmatrix} 0 & 0 & 4 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 6 & 2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 4 & 1 & 0 & 0 & 0 & 1 \\ -1 & -1 & -3 & -1 & 0 & 0 & 1 & -1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & 0 & -\frac{1}{4} & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

As this system is consistent, V $\subseteq$ S is true.
As dim(T) = 4 = number of columns in S, S is indeed a basis for V.

3.14.5 Finding a basis

Note: this section is actually in chapter 4 but is here for easier searching

Column Basis
To find a basis for the column space of a matrix A, find its RREF (matrix R). The basis for the column space are the vectors in matrix A, that correspond to the pivot columns in matrix R.

Row Basis
To find a basis for the row space of a matrix A, find its RREF (matrix R)
The basis for the row space are the nonzero rows in R.

Basis of a subspace defined on an equation
To find a basis of a subspace defined on an equation, e.g. $V = \{(a, b, c) \mid 2a+b = 0, c = 3\}$, express the system explicitly.
Put the equations in an augmented matrix and RREF to find relations between the columns. Express the solution set explicitly. This is the basis of the subspace.

Basis of solution set of linear system
If you're finding all the u such that Au=0, use this method
Find linear dependence in columns and express in terms of matrix: e.g.:

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Basis = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

3.15 Coordinates Relative to a basis

Let S = $\{u_1, u_2, .., u_k\}$ be a basis for V, a subspace of $\mathbb{R}^n$.
Then, the unique expression of a vector v can be given in terms of the basis S:

$$v = c_1u_1 + c_2u_2 + ... + c_ku_k$$

Where $c_1 - c_k$ are constants.
The vector in $\mathbb{R}^k$ defined by the coefficients of the equation above is called the coordinates of v relative to S:

$$[v]_s = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

3.15.1 Finding the coordinates relative to a basis

To find coefficients $c_1 - c_k$ such that the equation is fulfilled, is equivalent to:

$$\begin{pmatrix} u_1 & u_2 & ... & u_k \end{pmatrix} * \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = v$$

Putting this into an augmented matrix, solve for:

$$\begin{pmatrix} u_1 & u_2 & ... & u_k & | & v \end{pmatrix}$$

3.15.2 Properties of coordinates relative to a basis

Let V be a subspace, and B a basis of that subspace. | B| = k

1. Uniqueness: For any vectors u, v ∈ V, u=v iff $[u]_s = [v]_s$
2. Linear combinations: For any $v_1, v_2, .., v_m \in V$:

$$[c_1v_1 + c_2v_2 + ... + c_kv_k]_s = c_1[v_1]_s + c_2[v_2]_s + .. + c_k[v_k]_s$$

3. Linear Independence: $v_1, v_2, ..., v_m$ are linearly independent iff $[v_1]_B, [v_2]_B, ..., [v_m]_B$ are linearly independent in $\mathbb{R}^k$
4. Spanning $v_1, v_2, ...v_m$ spans V iff $[v_1]_B, [v_2]_B, ..., [v_m]_B$ spans $\mathbb{R}^k$

3.16 Dimensions

The dimension of a subspace V of $\mathbb{R}^n$, is defined to be the number of vectors in any basis of V.

The dimension of a solution space of a system is the number of non-pivot columns. Recall: Vectors of a general solution in a linear system form a basis for the solution space

3.16.1 Dimension Formula for Subspaces

Adding 2 subspaces together,

$$\begin{aligned}\dim(U_1 + U_2) &= |B_1 \cup B_2| \\ &= |B_1| + |B_2| - |B_1 \cap B_2| \\ &= |B_1| + |B_2| - |B_0| \\ &= \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).\end{aligned}$$

3.16.2 Spanning Set theorum

Let S = $u_1, u_2, ..., u_k$ be a subset of vectors in $\mathbb{R}^n$, and let V = span(S). Suppose V is not the zero space, V$\neq$0. Then there must be a subset of S that is a basis for V.

If S is linearly independent, the subset that fulfils this criteria is S itself.
If S is linearly dependent, then there exists a linearly independent subset that is a basis for V.

3.16.3 Linear Independence Theorum

Naturally from above. Let V be a subspace of $\mathbb{R}^n$ and S = $u_1, u_2, ..., u_k$ a linearly independent subset of V, S $\subseteq$ V. Then, there must be a set T s.t. S$\subseteq$T & T is a basis for V.

Case 1: Dimension of S is less than V => Add necessary vectors to form T such that dim(T) = dim(V)
Case 2: Dimension of S = V => dim(S) = dim(T) = dim(V)

3.17 Transition Matrix

Let S = $\{u_1, u_2, .., u_k\}$ and T = $\{v_1, v_2, .., v_k\}$. Given that S and T are bases of a subspace V, the transition matrix from T to S is the matrix P s.t.:

$$\mathbf{P}[w]_T = [w]_s$$

P *sends* the coordinates in T to coordinates in S.
P is given:

$$\mathbf{P} = ([v_1]_s \ [v_2]_s \ ... \ [v_k]_s)$$

This is equivalent to solving the augemented matrix:

$$\begin{pmatrix} u_1 & u_2 & ... & u_k & | & v_1 & v_2 & ... & v_k \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} I_k & \mathbf{P} \\ 0_{(n-k)*k} & 0_{(n-k)*k} \end{pmatrix}$$

3.17.1 Inverse of transition matrix

The inverse, $\mathbf{P}^{-1}$ is the transition matrix from S to T. However, you cannot assume P is invertible.

4 Subspaces associated to a matrix

4.1 Row Space

Simply put, the row space of a matrix A is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

$$\text{Row(A)} = \text{span}\{(a_{11}a_{12}...a_{1n}), (a_{21}a_{22}...a_{2n}), ..., (a_{n1}a_{n2}...a_{nn})\}$$

4.1.1 Row operations preserve row space

Row operations (done in RREF etc) **does not affect** the row space!
Suppose A and B are row-equivalent matrices. Then, Row(A) = Row(B)
While Row operations preserve linear relations between columns, they do not preserve linear relations between rows.

4.1.2 Properties of Row equivalent matrices

1. A is row-equivalent to U
2. A = PU, where P is some matrix representing the elementary row operations performed.
3. Rank is the same, rank(A) = rank(U)
4. Row relations preserved
5. Null space is reserved: null(A) = null(U)
6. Column space is not reserved: col(A) $\neq$ col(U)
   (a) Hence, Ax=c and Ux=c can have varying answers
   (b) However, the homogeneous linear system Ax=0 and Ux=0 will have the same answer.

4.1.3 Basis for Row space

A basis for a matrix A can be found by reducing it to RREF.
In the RREF, the nonzero rows of R form a basis for its row space.

4.2 Column space

The column space is what spanned by the columns.
Column space can be characterized by teh set of vectors **v** such that Ax = v is consistent, or the set of vectors v such that v = Au for some u. A vector is in the column space of A iff:

$$(u_1 u_2 ... u_k) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = c_1 u_1 + c_2 u_2 + ... + c_k u_k = \mathbf{v}$$

4.2.1 Row operations preserve linear relations between columns

If A = $(a_1 a_2 ... a_n)$ is row equivalent to B = $(b_1 b_2 ... b_n)$, then:

$$c_1 a_1 + c_2 a_2 + ... + c_n a_n = 0$$
$$\text{iff}$$
$$c_1 b_1 + c_2 b_2 + ... + c_n b_n = 0$$

The relations between columns are constant.

While Row operations preserve row space, they do not preserve column space.

4.3 Nullspace

The nullspace of a matrix A is the solution space to the system Ax=0 with coefficient matrix A. Denoted:

$$\text{Null(A)} = \{\ \mathbf{v} \in \mathbb{R}^n \mid Av = 0\}$$

Additionally,

$$\text{Null}(A^T A) = \text{Null(A)}$$

Intuition: The null spaces is the set of all vectors that get squooshed to 0 upon some transformation v. This corresponds to the solution set (free parameters) of the system - not all parameters get squooshed in dimension.
Note that this is opposite to the column space.
The nullity of A is the dimension of the nullspace of A.

4.3.1 Finding the span of the nullspace

Since this is simply a homogeneous system, RREF the matrix A and express the solution explicitly.

4.4 Rank

Rank is defined in many ways.

$$\text{Rank(A)} = \dim(\text{Col(A)}) = \dim(\text{Row(A)})$$

Rank(A) =
number of pivot columns in RREF of A, or
number of leading entries in RREF of A, or
number of nonzero rows in RREF of A

4.4.1 Properties of Rank

1. Rank is invariant under transpose
2. Rank(A) = 0 iff A = 0
3. A linear system Ax=b is consistent iff Rank(A) = Rank(A—b)
4. Rank(AB) $\leq$ min(Rank(A), Rank(B))
   From this we get: Col(AB) $\subseteq$ *Col(A)*
5. If A and R are row equivalent, Rank(A) = Rank(R)
6. Rank(A + B) $\leq$ rank(A) + Rank(B)

4.4.2 Rank-Nullity Theorum

If A is a m $\times$ n matrix,

$$\text{rank(A)} + \text{nullity(A)} = \text{n}$$

This is natural considering how rank counts the number of pivot columns, while nullspace the number of non-pivot columns.

4.5 Full Rank

A m$\times$n matrix **A** is full rank if its rank is equal to either the number of rows or columns.
A is full rank if and only if it's invertible.

4.6 Equivalent statements when Full Rank equals number of columns

Suppose A is a $m \times n$ matrix.

1. A is full rank, where rank is equal to number of columns, Rank(A) = m
2. The rows of A spans $\mathbb{R}^n$, Row(A) = $\mathbb{R}^n$
3. The columns of A are lienarly independent
4. The homogeneous system Ax=0 has only the trivial solution
5. Null(A) = 0
6. A^tA is an invertible matrix of order n
7. A has a left inverse

Pointing out #2, The columns of A are linearly independent, so there are n pivot rows. This means that there are (m-n) zero rows, and there are n nonzero rows.

4.7 Equivalent statements when Full Rank equals number of rows

1. A is full rank, where rank(A) = m = number of rows
2. The columns of A spans $\mathbf{R}^m$
3. The rows of A are linearly independent
4. The linear system Ax=b is consistent for every b $\in \mathbf{R}^m$
5. AA^t is an invertible matrix of order m.
6. A has a right inverse

5 Orthogonality, Projections, and Least Square Solution

5.1 Orthogonality

Two vectors u, v are orthogonal if

$$u \cdot v = 0$$

That is, they are perpendicular OR u or v are 0 vectors.

A set S = $\{v_1, v_2, ..., v_k\}$ of vectors in $\mathbb{R}^n$ are pairwise orthogonal if for every $i, j \in S$, $v_i * v_j = 0$.

The set S is orthonormal if for all $i, j \in S$

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

that is tranpose of S times S itself nets you the identity matrix,

$$S^T * S = I_n$$

S is orthogonal and all the vectors are unit vectors

5.1.1 Orthgonality under scaling

If you have an orthogonal set $\{p_1, p_2, p_3\}$, and you scale any of the vectors by ± 1, then they are still and orthogonal set.

However, any value other than ± 1 will ruin the orthogonality, unless they get scaled back.

5.1.2 Orthogonal to a subspace

Let V be a subspace of $\mathbb{R}^n$. A vector **n** is orthogonal to V i f for every v in V, $n \cdot v = 0$. That is, n is orthogonal to every vector in V.
Denoted,

$$n \perp V$$

5.1.3 Algortithm to check orthogonality

Given that V is spanned by a set of vectors $= \{u_1, u_2, ..., u_k\}$, to find orthogonality:

$$A^T w = 0$$

$$\begin{pmatrix} u_1^T & | & 0 \\ u_2^T & | & 0 \\ \vdots & \vdots & \vdots \\ u_k^T & | & 0 \end{pmatrix}$$

Where u_x^T are row vectors.

The nullspace of the system $A^T w = 0$ will be the set of vectors orthogonal to V.

Example: Getting the span of vectors orthogonal to the system
For V $= \{(w, x, y, z) \mid w - x + 2y + z = 0\}$

First, find a basis for V by expressing V explicitly. It's important you find the basis first

$$Basis = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Now, solve the system $B^T w = 0$. **You are transposing the basis!**

$$\begin{pmatrix} 1 & 1 & 0 & 0 & | & 0 \\ -2 & 0 & 1 & 0 & | & 0 \\ -1 & 0 & 0 & 1 & | & 0 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & -1 & | & 0 \\ 0 & 1 & 0 & 1 & | & 0 \\ 0 & 0 & 1 & -2 & | & 0 \end{pmatrix}$$

Now, get the nullspace:

$$w \perp V \Leftrightarrow w \in span$$

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

This is actually finding the orthogonal complement of V, which is the set of all vectors orthogonal to V. It is denoted:

$$V^{\perp} = Null(A^T)$$

5.1.4 Checking if vector is orthogonal:

To check if a vector is orthogonal to a given matrix, first find the Basis of the matrix.

$$\text{Basis} = \{u_1 \text{ ,..., } u_k \}$$

Put $u_1 ... u_k$ in a matrix U. To find if they are orthogonal, take:

$$U^T * v$$

This is equivalent to taking the dot product for each vector u_k in the basis. If they are 0, the vectors are orthogonal.

5.2 Orthogonal & Orthonormal Bases

Let V be a subspace of $\mathbb{R}^n$. A set S is an orthogonal basis for V if S is a basis for V and S is orthogonal.
If S is orthonormal, it is an orthonormal basis.
Weird placement? If S = $\{u_1, u_2, .., u_k\}$ is an orthogonal set of nonzero vectors, Then S is linearly independent.
Every orthonormal set is linearly independent

5.2.1 Coordinates relative to an Orthogonal Basis

S = $\{u_1, u_2, ..., u_k\}$ is an orthogonal basis for subspace V.
To express $v \in V$ relative to S it is given:

$$[v]_s = \begin{pmatrix} u_1 \cdot v / \|u_1\|^2 \\ u_2 \cdot v / \|u_2\|^2 \\ \vdots \\ u_k \cdot v / \|u_k\|^2 \end{pmatrix}$$

This is because when computing $u_i \cdot v$:
$u_i \cdot v = u_i \cdot (c_1 u_1 + c_2 u_2 + ...c_k u_k) = c_i u_i \cdot u_i$
This is because $u_i \cdot u_x$ where x≠i = 0, as they are orthogonal to each other.

5.2.2 Coordinates relative to an Orthonormal Basis

Consequently, if S is orthonormal, the coefficients to get v is simply the dot product:

$$[v]_s = \begin{pmatrix} u_1 \cdot v \\ u_2 \cdot v \\ \vdots \\ u_k \cdot v \end{pmatrix}$$

5.3 Orthogonal Matrices

A $\underline{n \times n}$ square matrix **A** is orthogonal if $A^T = A^{-1}$. Or, $A^T A = I = AA^T$

Equivalent Statement

1. A is an orthogonal matrix
2. The columns of A form an orthonormal basis for $\mathbb{R}^n$
3. the rows of A form an orthonormal basis for $\mathbb{R}^n$

Points #2 and #3:

$$a_i^T a_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

From this, we can see that all vectors are $\perp$ to each other, and since the diagonal is 1, the vectors are unit length.

5.4 Orthogonal Projection

We use the orthogonal projection to project a vector that is not in the subspace into the subspace.
The projection is given:

$$w_p = \frac{w \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{w \cdot u_2}{u_2 \cdot u_2} u_2 + ... + \frac{w \cdot u_k}{u_k \cdot u_k} u_k$$

Where u_n are the vectors forming the basis.
You can also use this formula:

$$w_p = B(B^T B)^- 1 B^T w$$

where B is the basis of the space.
The projection of a vector onto the space, is the vector in V that is closest to w.

Orthonormal projection Similarly,

$$w_p = (w \cdot u_1)u_1 + (w \cdot u_2)u_2 + ... + (w \cdot u_k)u_k$$

5.4.1 Gram-Schmidt Process

The Gram-Schmidt process converts a normal basis for a subspace V into an orthonormal basis. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a linearly independent set. Let

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{v}_1 \cdot \mathbf{u}_2}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{v}_1 \cdot \mathbf{u}_3}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_3}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_k &= \mathbf{u}_k - \frac{\mathbf{v}_1 \cdot \mathbf{u}_k}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_k}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \cdots - \frac{\mathbf{v}_{k-1} \cdot \mathbf{u}_k}{\|\mathbf{v}_{k-1}\|^2}\mathbf{v}_{k-1}.
\end{aligned}$$

Note: $\|\mathbf{v}_k\|^2$ is equivalent to $v_k \cdot v_k$

Then $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is an orthogonal set (of nonzero vectors), and hence,

$$\left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \dots, \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} \right\}$$

MATLAB tip

Yes i spelt it wrong. Use symbolic toolbox to get exact value, sym(A)

```
function Q = gramschidt(A)
n = size(A,1);
for j=1:n
    v=A(:,j);
    for i=1:j-1
        R(i,j)=Q(:,i)'*A(:,j);
        v=v-R(i,j)*Q(:,i);
    end
    R(j,j)=norm(v);
    Q(:,j)=v/R(j,j);
end
Q;
end
```

Shout out MIT

5.5 QR Factorization

QR factorization decomposes a linearly independent (rank(A) = n) $m \times n$ matrix A into 2 matrices:

$$A = QR$$

where Q is a $m \times n$ matrix s.t. $Q^T Q = I_n$ and R is an invertible upper triangular matrix with positive diagonal entries. Q is an **orthonormal basis** for A.

5.5.1 Algorithm to QR Factorize

1. Perform Gram-Schmidt process on columns of A to obtain an orthonormal set $\{q_1, q_2, ..., q_n\}$
2. Set Q = $(q_1 q_2 ... q_n)$
3. Compute R = $Q^T A$

Note: In computing R, we are essentially computing the projection of the vectors of A into the orthonormal basis Q.
R is the coordinates of A relative to the orthonormal basis Q.

Q is orthonormal
Each vector of Q is orthogonal to each other, and they must be normalized.

5.5.2 Manipulating QR Factorizations

When manipulating QR factorizations, keep in mind that $Q'T * Q = I_n$, because Q is orthonormal. Using this, you can manipulate equations such that:

$$A^T * A = (QR)^T * QR = R^T * Q^T * Q * R = \mathbf{R^T} * \mathbf{R}$$

5.5.3 Eligibility of QR factorization

1. A is QR factorizable
2. A has full column rank
3. For a $m \times n$ matrix, $m \geq n$

5.6 Least Square Approximations

Let A be a $m \times n$ matrix and b a vector in $\mathbb{R}^m$. A vector u in $\mathbb{R}^n$ is a least square solution to Ax=b if and only if Au is the projection of b onto the column space of A.

Let A be a m × n matrix and b a vector in $\mathbb{R}^m$. A vector u in $\mathbb{R}^n$ is a least square solution to Ax = b if and only if u is a solution to $A^T Ax = A^T b$.

Note that the least square solution is **not** unique. However, its projection, Au is unique.

5.6.1 Algorithm to find Least Square Approximations

1. Compute $A^T A$ and $A^T b$
2. Find the general solution of $(A^T A|A^T b)$
3. The general solution is the values of u, or the <u>least square solution</u>
4. Multiply and u to Au to find the least square projection.

nullspace of that matrix is the values that will squoosh the vector b onto the plane.

5.6.2 Shortcut: least squares to fix over-determined systems

Let's say you have a linear system $a_1x_1 + a_2x_2 + a_3x_3 = b$. You want to find the values a_1, a_2, a_3 such that you can find the solution to the system and do some math on it.
But, the RREF of the linear system is **inconsistent**. Then, to find the best coefficients a_1, a_2, a_3 s.t. the system "works":

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (A^T * A)^{-1} * A^T * b$$

This will give you the best coefficients such that you can plug in new x-values.

5.6.3 Manipulation of Least square approximations

If you want to find some v such that Av=0, or v is in the nullspace of A, and for some reason you don't have A but you do have 2 separate least square solutions for A, you're in the right place.

$$A^T A x_1 = A^T b$$
and
$$A^T A x_2 = A^T b$$
So,
$$A^T A(x_1 - x_2) = A^T b - A^T b$$

$x_1 - x_2$ is in the nullspace of $A^T A$, which is equivalent to nullspace(A).

5.6.4 Unique Least Square Solutions

If A has **full column rank**, then $A^T A$ is invertible, and the least square solution is unique.
If A is invertible, then finding the least square solution is equal to just solving Ax = b.

6 Eigenanalysis

6.1 Eigenvalues and Vectors

For linear transformations, some particular vectors get amplified in magnitude while retaining direction. This is the basis of eigenanalysis.

For a square matrix A of order n, $v \neq 0$

$$\text{Av} = \lambda v$$

The real number λ is an eigenvalue of the transformation A. v is the eigenvector associated to λ

6.1.1 Characteristic Polynomial

The characteristic polynomial finds all the eigenvectors of a transformation A. It is given:

$$\text{det(xI-A)}$$

This is because λ is an eigenvalue iff:

$$(\lambda I - A)v = 0$$

v is nonzero, hence, v is a nontrivial solution to the homogeneous system.

The, λ is an eigenvalue of A iff it is a root of the characteristic polynomial

MATLAB tip

Calculate the characteristic polynomial with the following:

```
% Set symbols
syms x
% create the augmented matrix
A = [0 2 3 8; 4 2 2 7; a 7 3 1;]
% get size to make order n identity matrix
% !! assumes square matrix !!
[rows,columns] = size(A)
% find characteristic polynomial
factor(det(x*eye(rows) - A))
```

With minimal tweaking, you can use this to calculate the eigenspace as well.

6.1.2 Algebraic Multiplicity

Provided the determinant of the characteristic polynomial split into linear factors:

$$det(xI - A) = (x - \lambda_1)^{n_1}(x - \lambda_2)^{n_2}....$$

The Algebraic Multiplicity is the highest power of all the roots:

$$\text{Max}(n_1, n_2, ..., n_x)$$

It is also how many times the eigenvalue appears in the root.

Algebraic multiplicity is **always** more than or equal to geometric multiplicity.

6.1.3 Eigenvalues of Triangular Matrices

The eigenvalues of a triangular matrix are the diagonal entries. The algebraic multiplicity of the eigenvalue is the number of times it appears in the diagonal.

6.1.4 Eigenspace

The eigenvectors associate to A is given:

$$(\lambda I - A)x = 0$$

This system is homogeneous, and hence the set of solutions is a subspace. This is the eigenspace of A associated to eigenvalue λ
It is given by:

$$\text{Null}(\lambda I - A)$$

6.1.5 Eigenvalue 0

The eigenvalue 0 sends all associated eigenvectors to 0 upon some transformation. This is a particularly important for several things, like least square projection where the solution is based on the nullspace of A.
Equivalent statements of eigenvector 0

1. The matrix has an eigenvalue 0
2. The matrix is not invertible
3. Eigenvectors associated with 0 are the **nonzero** vectors associated with A$v = 0$; the nullspace of A.
4. The matrix is nilpotent if 0 is the **only** eigenvalue.
5. The nullity of A is given by the geometric multiplicity of eigenvalue 0

6.1.6 Geometric Multiplicity

The geometric multiplicity of an eigenvalue λ is the dimension of its eigenspace:

$$\text{nullity}(\lambda I - A)$$

6.1.7 Independence of Eigenspaces

Vectors associated to distinct eigenvalues of A are linearly independent

6.2 Diagonalization

A square matrix A is diagonalizable if there exists an invertible matrix P s.t.

$$P^{-1}AP = D$$

where D is a diagonal matrix.

$$\text{A} = PDP^{-1}$$

A consequence of this is that you can define a matrix uniquely by knowing its eigenspace and eigenvectors.

Powers of Diagonalizable matrix By doing this, it makes it much easier to calculate powers on the matrix A.

$$A^m = PD^mP^{-1}$$

A is diagonalizable iff:

$$P = (u_1 u_2 ... u_n),\ D = \begin{pmatrix} d_1 & 0 & ... & 0 \\ 0 & d_2 & ... & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & ... & d_n \end{pmatrix}$$

where d_i is the eigenvalue associated to eigenvector u_i. d_i *may* not be distinct.

Thus, there must be n independent eigenvectors for P. If there are n distinct eigenvalues, it guarantees n distinct eigenvectors, and thus would be diagonalizable.

6.2.1 Equivalent statements for Diagonalizability

1. A is diagonalizable
2. There exists a basis $\{u_1, u_2, ..., u_n\}$ of $\mathbb{R}^n$ of eigenvectors of A.
3. The characteristic polynomial of A splits into linear factors
4. The geometric multiplicity for each eigenvalue is equal to the algebraic multiplicity

If any of these are not met, A is not diagonalizable.

6.2.2 Example of Diagonalization

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

Find determinant of characteristic polynomial (code above): $(x-2)^2(x-4)$
Find a basis for the eigenspaces:

$$2I - A = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Basis = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Note that the gemoetric multiplicity, the number of nonpivots for the matrix, is **2**. This matches the algebraic multiplicity.

$$4I - A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, Basis = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Again, note that the geometric multiplicity matches the algebraic multiplicity, 1.
Finally, A is diagonalizable with:

$$P = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

6.3 Orthogonally Diagonalizable

A is orthogonally diagonalizable if

$$A = PDP^T$$

For some orthogonal matrix P and diagonal matrix D.

P is orthogonal iff its columns $u_1, u_2, ..., u_n$ is an orthonormal basis for $\mathbb{R}^n$

Spectral Theorum A is orthogonally diagonalizble if and only if A is symmetric.

6.3.1 Equivalent Statments for Orthogonally Diagonalizable

1. A is orthogonally diagonalizable.
2. There exists an orthonormal basis $\{u_1, u_2, .., u_n\}$ of $\mathbb{R}^n$ of eigenvectors of A. (They don't have to already be orthonormal - we can normalize it)
3. A is a symmetric matrix.

6.3.2 Independence of Eigenspaces of symmetric matrix

Vectors associated to distinct eigenvalues of symmetric matrix A are linearly independent

6.3.3 Algorithm to orthogonally diagonalize

As for the normal one, compute the characteristic polynomials and find the basis for the solution set of each $(\lambda I - A)x = 0$.
Now, apply Gram-Schmidt process on **each** basis (reiterate: one-by-one for each basis) to obtain an orthonormal basis for $\mathbb{R}^n$
Set P to the orthonormal bases and D to the diagonal as per usual.

6.4 Markov Chains

1. A vector v with nonnegative coordinates that add up to 1 is called a probability vector
2. A stochastic matrix is a square matrix whose columns are probability vectors.
3. A markov chain is a sequence of probability vectors x_0, x_2, x_k together with a stochastic matrix P such that:

$$x_1 = Px_0, \quad x_2 = Px_1, \quad x_k = Px_{k-1}...$$

4. x_k is also called the state vector

6.4.1 Example

Spongebobby is an avid anime figurine collector. However, he only purchases 3 kind of anime figurines: Naruto, Dragonball, and, Hunter X Hunter (HXH). If Spongebobby buys a Naruto figurine, there is a 20% chance he buys it again, and a 40% chance he buys a DragonBall Figurine next. If he buys a Dragonball figurine, there's a 30% chance he buys a Naruto one next and a 70% chance he buys HxH. If he buys a HxH figurine, there's a 10% chance he buys Naruto, 60% chance DB.

There are 3 state vectors: He buys Naruto (e_1), Dragonball (e_2), and HxH(e_3)

$$Pe_1 = \begin{pmatrix} 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \quad Pe_2 = \begin{pmatrix} 0.3 \\ 0 \\ 0.7 \end{pmatrix} \quad Pe_3 = \begin{pmatrix} 0.1 \\ 0.6 \\ 0.3 \end{pmatrix} \Rightarrow P = \begin{pmatrix} 0.2 & 0.3 & 0.1 \\ 0.4 & 0 & 0.6 \\ 0.4 & 0.7 & 0.3 \end{pmatrix}$$

The state vector after n days will be $x_n = P^n x_0$

Using this, we can calculate what spongebobby will probably buy after n days, given the starting vector x_0. x_0 represents what spongebobby buys on the first day.

6.4.2 Equilibrium Vector

The equilibrium vector for a stochastic matrix P is a probability vector that is an eigenvector associated to eigenvalue 1.

It represents the long-term distriution.

Finding the equilibrium vector

1. Find any eigenvector associated to eigenvalue 1. That is, find the basis "B" of the solution space of the homogeenous linear system (I-P)x = 0. Let the vector spanning B be equal to U, {U}.
2. Let s = sum of coordinates of the vector U.
3. Multiply s^{-1} to the vector U.

6.5 Single Value Decomposition

A factorization of A ($m \times n$) into:

$$\mathrm{A} = \mathrm{U}\Sigma V^T$$

where U is an order **m** orthogonal matrix, V an order **n** orthogonal matrix, and the matrix Σ has the form:

$$\Sigma = \begin{pmatrix} D & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$$

for some diagonal matrix D of order r, where r $\leq$ min{m,n}. The size of sigma is $m \times n$
is a SINGLE VALUE DECOMPOSITION.

6.5.1 Ranks of SVD

Rank of U corresponds to m.
Rank of sigma corresponds to Rank(A).
Rank of V corresponds to n

6.5.2 Singular Values

If A is a $m \times n$ matrix, then $A^T A$ is an order n symmetric matrix, and we may orthogonally diagonalize it. Then, the square root of eigenvalues of $A^T A$ are the singular values of A, $\sqrt{\lambda_x}$
These values populate the diagonal of matrix Σ

6.5.3 Algorithm to SVD

1. Get Eigenvalues of $A^T A$ and Singular Values of A.
2. The singular values form Σ
3. Using the Eigenvalues of $A^T A$, find the basis of each eigenspace. Take the unit vector associated to each space - normalize if required.
4. The unit vectors form V.
5. Set u_x = Singular value
6. For each u_x, take $u^x A v_x$, where v_x are the unit vectors identified before.
7. These form a column of U.
8. If U is not of correct size, extend to find an orthonormal basis of the right size.
   (a) To do this, solve $(u_1 ... u_r)^T x = 0$
9. Perform gram-schmidt if not an orthonormal basis

6.5.4 Example of SVD

Let $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \Rightarrow \det(xI - A^T A)$ = (x)(x-18)

Singular values = $\sqrt{0} = 0$ and $\sqrt{18} = 3\sqrt{2}$

$$\Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$18I - A = \begin{pmatrix} 9 & 9 \\ 9 & 9 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad Basis = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\} \xrightarrow{Normalize} \begin{pmatrix} -1\sqrt{2} \\ 1\sqrt{2} \end{pmatrix}$$

$$0I - A = \begin{pmatrix} -9 & 9 \\ 9 & -9 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad Basis = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \xrightarrow{Normalize} \begin{pmatrix} 1\sqrt{2} \\ 1\sqrt{2} \end{pmatrix}$$

Thus, V $= \begin{pmatrix} -1\sqrt{2} & 1\sqrt{2} \\ 1\sqrt{2} & 1\sqrt{2} \end{pmatrix}$

Then, $u_1 = (3\sqrt{2})^{-1}$ * A * $v_1 = \begin{pmatrix} -1/3 \\ 2/3 \\ -2/3 \end{pmatrix}$

However, U needs to be an order 3 orthonormal basis. So, we extend it General solution for U_1=0:

$$s \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

We find two independent vectors from this general solution:

$$(1) \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + (0) \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

$$(0) \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + (1) \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Perform gram-schmidt process to obtain:

$$\text{U} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{pmatrix}$$

7 Linear transformations

7.1 What are they

A mapping/function T: $R^n \to R^m$ is a linear transformation if for all u,v in $\mathbb{R}^n$ and scalars α, β

$$\mathrm{T}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T((v))$$

Additionally, it **must** map the 0 vector in the domain to the 0 vector in the codomain
$\mathbb{R}^n$ is called the domain of the mapping, and $\mathbb{R}^m$ is the codomain.

7.2 Standard Matrix

The linear transformation T: $\mathbb{R}^n \to \mathbb{R}^m$ can be represented by a unique $m \times n$ matrix A such that:

$$\mathrm{T(u) = Au}$$

The matrix A is given by:

$$\mathrm{A} = (\mathrm{T}(e_1)\ \mathrm{T}(e_2)\ ...\ \mathrm{T}(e_n))$$

where E $= \{e_1, e_2, ..., e_n\}$ is the standard basis for $\mathbb{R}^n$

7.3 Standard Matrix w.r.t. Basis

If S $= \{u_1, u_2, ...u_n\}$ is a basis for $\mathbb{R}^n$, the representation of T with respect to basis S, $[T]_s$, is defined:

$$[T]_s = (T(u_1)\ T(u_2)\ ...\ T(u_n))$$

Then, A $= [T]_E$ - perform change of basis to standard basis E.

7.4 Algorithm to find standard matrix

1. Ensure that $\{v_1, v_2, ..., v_n\}$ form a basis for the subspace $\mathbb{R}^n$
2. Write $[T]_s = (T(v_1) \quad T(v_2)...T(v_n))$
3. If $v_1 - v_n$ is not the standard basis (identity matrix), get the transition matrix to the standard basis:

   Transition matrix P = rref($v_1\ v_2\ ...\ v_n$ — I_n) = $(v_1 \quad v_2 \quad ... \quad v_n)^{-1}$
4. Then, A $= [T]_s * P$

7.5 Range and Kernel of Linear Transformation

For all, assume T: $\mathbb{R}^n \to \mathbb{R}^m$ and A is the standard matrix of T

7.5.1 Range

$$\text{Range} = \{\ v \in \mathbb{R}^m \mid v = T(u) \text{ for some u} \in \mathbb{R}^n\}$$

This is the Col(A)

7.5.2 Rank

Rank of T is:

$$\text{Rank(T)} = \text{dim(Range(T))}$$

This is also the rank(A)

7.5.3 Kernel

Kernel of T is:

$$\text{kernel(T)} = \{\ \text{u} \in \mathbb{R}^n \mid \text{T(u)} = 0\ \}$$

Of the standard matrix A, it is the nullspace.

7.5.4 Nullity

Nullity of T is:

$$\text{nullity(T)} = \text{dim(ker(T))}$$

It is also nullity(A)

7.5.5 Injectivity

T is injective / one-to-one if for every vector in range of T, there is a unique **u** such that T(u)=v

T is injective if and only if $\mathbf{ker(T) = 0}$

Additionally, T is injective iff full rank equals number of columns

7.5.6 Surjectivity

T is surjective / onto if for every v in codomain, there exists a u in the domain s.t. T(u) = v

T is surjective iff full rank equals number of rows

7.5.7 Bijectivity

Not formally defined, but if T is both surjective and injective (bijective), it is naturally a square matrix and invertible.