

1. Scalar: 3, 40, -4

2. Vector

- definition: an ordered list of scalars
- categories: row vector, column vector
- special case: zero vector
- operations:
 - equal: same size + same corresponding elements
 - in \mathbb{R}^n (nth real dimension)
 - addition
 - scalar multiplication
 - **dot product:**

$$\begin{bmatrix} 1 \\ 5 \\ 8 \end{bmatrix} \cdot \begin{bmatrix} 12 \\ 3 \\ 6 \end{bmatrix} = (1 \times 12) + (5 \times 3) + (8 \times 6) = 12 + 15 + 48 = 75$$

- Draw a 2-dim vector in \mathbb{R}^2 and its addition - the parallelogram rule

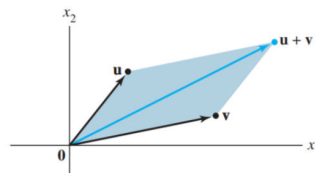


FIGURE 3 The parallelogram rule.

3. Matrix

- definition:
 - a rectangular array of scalars.
 - concatenate row vectors vertically
 - concatenate column vectors horizontally
- indexing: are given by the row and column number of the element.

$$\begin{bmatrix} 2_{1,1} & 9_{1,2} & 67_{1,3} & 78_{1,4} \\ -5_{2,1} & 64_{2,2} & 21_{2,3} & 12_{2,4} \\ 8_{3,1} & 19_{3,2} & 32_{3,3} & 77_{3,4} \\ 83_{4,1} & 17_{4,2} & 2_{4,3} & -4_{4,4} \\ 28_{5,1} & 11_{5,2} & 54_{5,3} & 19_{5,4} \end{bmatrix}$$

- main diagonal of the matrix
- special case: square matrix, **identity matrix**
- Operations:
 - addition
 - scalar multiplication
 - **transpose:** flip row and column A^T
 - **Matrix multiplication:**
 - constrain: inner dimensions must match

$$\begin{array}{c}
 \text{Size of result = outer dimensions} \\
 m \times n \quad \text{Inner dimensions must match} \quad p \times q
 \end{array}$$

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{1,m} & \cdots & a_{m,n} \end{bmatrix} \quad B = \begin{bmatrix} b_{1,1} & \cdots & b_{1,q} \\ \vdots & \ddots & \vdots \\ b_{1,p} & \cdots & b_{p,q} \end{bmatrix}$$

- Matrix power: only possible for square matrices

$$\begin{aligned}x_1 + 4x_2 + 12x_3 &= 18 \\7x_1 + 3x_3 &= 21 \\6x_1 - 10x_2 + 15x_3 &= 45\end{aligned}$$

$$\begin{matrix} & A & & x & = & b \\ \begin{bmatrix} 1 & 4 & 12 \\ 7 & 0 & 3 \\ 6 & -10 & 15 \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} & = & \begin{bmatrix} 18 \\ 21 \\ 45 \end{bmatrix} \end{matrix}$$

Fig. 1: linear system of equations

1. Linear System of Equations

- Matrix form: $Ax=b$
- Augmented matrix: $[A \ b]$

2. How to solve the problem?

- Row Reduction
 - principle: **eliminate** variables in other equations to obtain a simplified **equivalent system of equations**
 - three elementary row operations:
 - **swap** any two rows
 - multiply any row by a non-zero scalar
 - add any scalar multiple of a row to another row
 - **Row Echelon Form:**
 - definition:
 - any rows of all zeros appear at the bottom of the matrix
 - the first non-zero entry (**leading entry, pivot entry**) in each row, (reading from the left), occurs to the right of the one in the row above it.
 - all entries in a column **below** a leading entry (pivot entry) are **zeros**.
 - properties:
 - not unique
 - if the augmented matrix is in row echelon form, the system can be solved using back-substitution

$$\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

Fig.2: row echelon form

• **Reduced Row Echelon Form**

- definition:
 - row echelon form
 - every pivot entry is **one**
 - all entries in a column **above** each pivot are **zeros**
- properties:
 - **unique** for every matrix
 - if the augmented matrix is in reduced row echelon form, then the system can be solved easily by inspection.

$$\begin{bmatrix} 1 & 0 & 0 & -10 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

Fig.3: reduced row echelon form

3. How many solutions do we have?

- **Free Variables & Basic Variables**
 - variables corresponding to pivot columns are called basic variables
 - otherwise, free variables

$$\left[\begin{array}{cccc|c} 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$x_1 \quad x_2 \quad x_3 \quad x_4$

- x_2 and x_4 are basic variables
- x_1 and x_3 are free variables

$$\begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Fig.4: inconsistent system

- inconsistent system: **no solution**
 - a system of linear equations is inconsistent **if and only if** a row echelon form of its augmented matrix has a pivot in the **right most column**
- consistent system:
 - **unique** solution: **no free** variables
 - **infinite** solution: has free variables

(continue from lecture#2)

4. For linear system of equations, which has infinite solution, how do we **parameterize** all solutions?

- principle: write all variables in terms of **free** variables
- solution set = **particular vector** + **linear combination of spanning vectors**
- there is only one spanning vector for each free variable

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 5 \\ 0 \\ 6 \end{bmatrix}}_p + x_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{v_1} + x_3 \underbrace{\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}}_{v_2}$$

5. **linear combination**:

- definition:
 - A linear combination of vectors is a **weighted sum of** a set of n vectors: $[v_1, v_2, v_3, \dots, v_n]$, written as: $c_1v_1 + c_2v_2 + c_3v_3 + \dots + c_nv_n$, where the c values can be any scalar
- example:

$$c_1 \begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 4 \\ -7 \end{bmatrix}$$

6. **span**:

- definition:
 - the span of vectors is **the set of all their linear combinations**.
 - those original vector set are called the **spanning vectors**
- shape:
 - the set of all points (vectors) we can reach through scaling & addition is the span

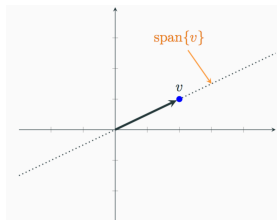


Fig 1: span of a nonzero vector - **straight line** passing through v and 0

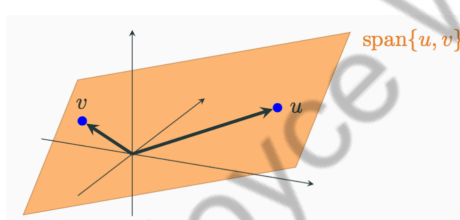


Fig 2: span of two **independent** vectors - **plane** passing through u , v and 0

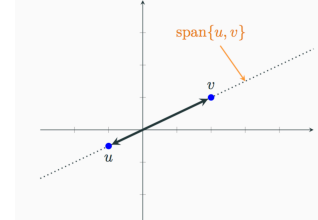
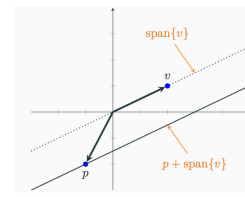
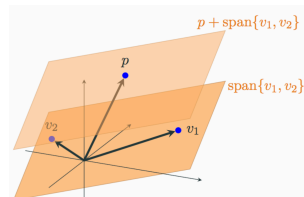


Fig 3: span of two **colinear** vectors - **straight line** passing through u , v and 0

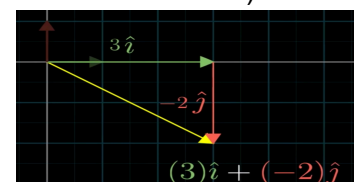
7. **Homogeneous Equation** $Ax=0$

- **trivial** solution: the zero vector
- What is the correlation between solution sets of $Ax=b$ (consistent case) and $Ax=0$?
 - property#1: have the same number of solutions
 - reason: #free-variables in rref (reduced row echelon form) are the same
 - property#2: if p is the particular solution of $Ax=b$, v is solution set of $Ax=0 \Rightarrow w=p+v$ is the solution set for $Ax=b$
 - property#3: if $w=p+v$ is the solution set of $Ax=b$, v is the span part $\Rightarrow v$ is the solution set of $Ax=0$
- Geometrically: The solution set of $Ax=b$ is a line/plane through p parallel to the solution set of $Ax=0$



8. Geometric understanding

- vector addition: combining two **movements**, represented by each vector
- Anytime we describe vector numerically, it depends on an implicit choice of what **basis vectors** we are using.
- vector is a linear combination of the coordinate system's basis vectors (the value of each entry in the vector corresponds to the coefficient/scalar of each basis vector)



1. Linear Independence

- definition:
 - The vectors in a set $T = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ are said to be *linearly independent* if the equation: $a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_n \vec{v}_n = \vec{0}$, can only be satisfied by $a_i = 0$ for $i = 1, \dots, n$
 - This implies that **no** vector in the set can be represented as a linear combination of the remaining vectors in the set.
- special case: v_1 and v_2 are called **colinear** \Leftrightarrow one of the vectors is a scalar multiple of the other.
- row reduction** is the easiest way to tell which vectors in a set are linearly dependent or independent

2. Revisit $Ax=b$

- if the vector b is inside the span of the columns of A (the matrix equation has a solution), then b can be written as a *linear combination of the columns of A* .

$$\begin{array}{c} A \\ \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \end{array} \begin{array}{c} x \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{array} = \begin{array}{c} b \\ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \end{array}$$

$$x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{bmatrix} + x_3 \begin{bmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \end{bmatrix} = b$$

3. The **essence** of matrix-vector multiplication ($Ax=b$)

- Applying a certain linear transformation to the vector

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$$

- this linear transformation is completely **determined** by where it takes the **basis vectors** of the space.
- $[(a,c)$ and $(b,d)]$ are the two basis vectors $[(1,0)$ and $(0,1)]$ moved to after the transformation

- Identity** transformation:

- the transformation does nothing, the i -hat and j -hat remains unmoved.

4. Linear transformation

- Numerically: formal linearity properties:

- $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- $cT(\mathbf{u}) = T(c\mathbf{u})$

- Geometrically:

- origin remain fixed
- lines remain lines without getting curved \Rightarrow keeping grid lines parallel and equally space.

- examples:

- scaling, rotation, reflection, shear (scale the vector differently in different dimensions), projection

- what is invertible transformation?

- for every $u \in \mathbb{R}^m$ there is one and only one $v \in \mathbb{R}^n$ such that $u = T(v)$

- Corresponding inverse transformation takes u to v , $T^{-1}: \mathbb{R}^m \rightarrow \mathbb{R}^n$

5. Matrix Inverse

- A^{-1} is the **unique** transformation with the property that if you first apply A , then followed it with the transformation A^{-1} . You end up back where you started.

- $AA^{-1} = A^{-1}A = I$

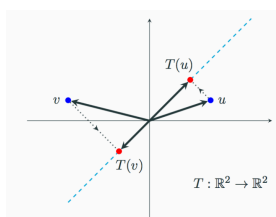


Fig1: Projection.
Not invertible

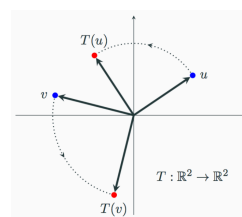


Fig2: Rotation.
Invertible

1. Invertible Transformation: onto & one-to-one

- definition: for **every** $u \in R^m$ there is **one and only one** $v \in R^n$ such that $u = T(v)$
- onto \Leftrightarrow every:
 - every vector b in the transformed space has one/multiple correspondence in the original space
 - $Ax=b$ has solution for every $b \Leftrightarrow Ax=b$ is consistent for every b
 - A has a pivot **in every row**
- one-to-one \Leftrightarrow one and only one: if
 - if b has correspondent mapping in the original space, then that mapping is unique.
 - all variables in the system are basic variables, no free variables
 - A has a pivot **in each column**
- invertible/isomorphism
 - one-to-one & onto

2. Determinant

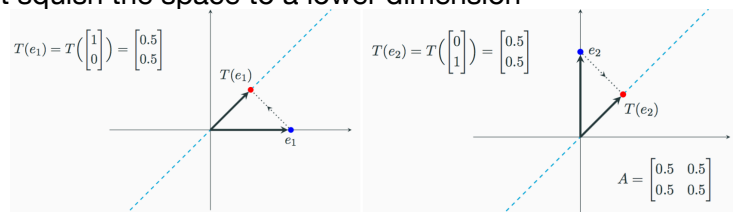
- definition geometrically:
 - tells how much a transformation scales area
 - the factor by which a linear transformation changes any area
 - the **sign** represents the orientation of the space. If determinant is **negative**:
 - space gets flipped over
 - the orientation of the space has been inverted
 - in 2D case, \hat{j} is to the right of \hat{i}
- compute:
 - 2x2 case $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
- properties:
 - if A and B are $n \times n$ matrices, then $\det(AB) = (\det A)(\det B)$

3. Revisit $Ax=b$: $\det(A)$ & A^{-1}

- if $\det(A) = 0$:
 - the transformation A squishes space into a smaller dimension (think about projection: plane \rightarrow line, *figure below*)
 - A^{-1} doesn't exist!
 - you cannot un-squish a line to turn it into a plane (not one-to-one)
 - solution might exist
 - if we are lucky enough, that the vector b lives on that transformed line/plane
- if $\det(A) \neq 0$:
 - space doesn't get squished into a lower dimension
 - can always find the **unique** vector x lands on b after the transformation A
 - A^{-1} exist! (one-to-one, onto)
 - solution exist: $x = A^{-1}b$, where $AA^{-1} = I$
- Note: A^{-1} exist $\Rightarrow Ax = b$ have unique solution; A is not invertible $\neq Ax = b$ has no solution

4. If A is invertible:

- A is a **square matrix**.
- A has a pivot in every row $\Leftrightarrow A$ has a pivot in every column
- the columns of A are linearly independent
- each of the column in A cannot be written as a linear combination of the other columns
- the transformation A represents doesn't squish the space to a lower dimension
- determinate of A is not ZERO



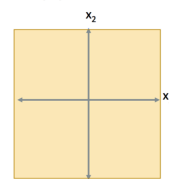
1. subspace, column space, null space and rank.

- subspace: any region within a space R^n that is expressible as a span
 - contains the zero vector
 - is the span of the collection of all vectors
- column space:
 - the **span** of the columns of your matrix, where columns are the basis vectors of the space
- rank**:
 - the #dimensions in the output of a transformation / in the column space
 - full rank**: the column space equals #columns, is as high as it can be.
 - columns in standard matrix A are linearly independent
 - A has a pivot in each column
 - for a square matrix A, A is full rank \Leftrightarrow A is invertible (geometrically, the space doesn't get squished after the transformation)

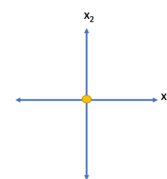
$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -2 \\ 0 & 0 & -4 \end{bmatrix}$$
- null space:
 - the set of vectors that lands on the origin, after the transformation
 - $Ax=0$, null space gives us all possible solutions to the equation
- nullity:
 - the dimension of the null space of a matrix
- if matrix A has n columns: $rank(A) + nullity(A) = n$

• Example: Graph the columns space of null space of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Col(A) = all of R^2



Null(A) = the zero vector



2. How to compute A-inverse?

$$[A | I] \xrightarrow{\text{rref}} [I | A^{-1}]$$

3. revisit $Ax=b$, the choice of basis

$$Ax=b \Rightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

• x is the coefficients:

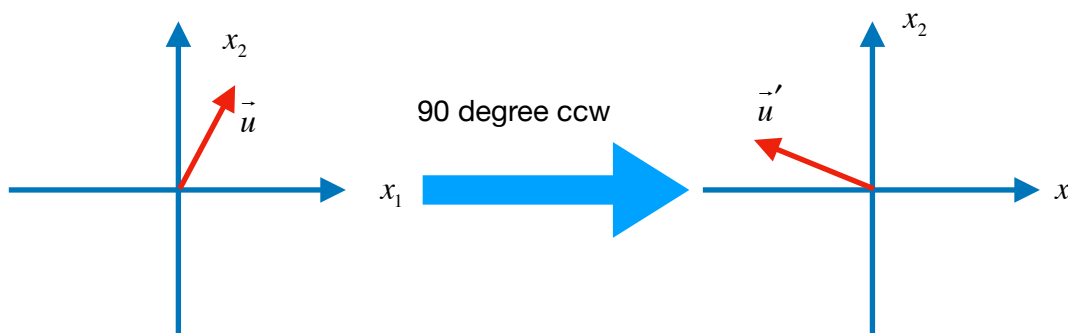
- the linear combination / coordinate of the original vector \vec{u} , based on original basis \hat{i}, \hat{j}

- the linear combination / coordinate of the transformed vector \vec{u}' , based on transformed basis \hat{i}', \hat{j}'

- the original vector \vec{u} and transformed vector \vec{u}' share the same linear combination / coordinate out of two different basis. (as a result of the property: the grid lines remain parallel and evenly spaced)

- b is the linear combination / coordinate of the transformed vector \vec{u}' , based on **original basis** \hat{i}, \hat{j}

•



1. Eigenvector & Eigenvalue

- **Eigenvector** is a special vector that after linear transformation A, it turns into scalar multiples of itself.
 - they are the ones that stay on their own span during a transformation.
- **Eigenvalue** is the factor by which it is stretched or squashed during the transformation.
- $Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$, x is a **nonzero vector (eigenvector)** and λ is a scalar (eigenvalue)
 - reason: identity matrix multiple any matrix won't change that matrix
- **Characteristic equation:**
 - $\det(A - \lambda I) = 0$
- How to compute eigenvalue?
 - homogeneous system $(A - \lambda I)x = 0$ has **nontrivial/nonzero** solution for particular λ
 - matrix $A - \lambda I$ is not invertible $\Leftrightarrow \det(A - \lambda I) = 0$
 - get λ
- How to compute eigenvector with respect to specific eigenvalue λ ?
 - eigenvector is the solution set of the homogeneous system $(A - \lambda I)x = 0$
 - rref of $A - \lambda I$; write all variables in terms of free variables; parametric form.

2. Basis

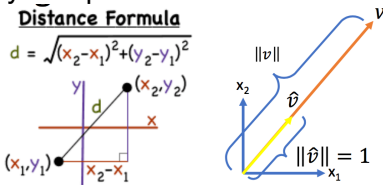
- a basis for a subspace is a linearly independent set whose span is the subspace
- **standard** basis: identity matrix
- a basis for the column space A is formed from the linearly independent columns of A

3. magnitude of a vector:

- the vector's length / the distance from tail to head
- $\|u\| = \sqrt{u \cdot u}$

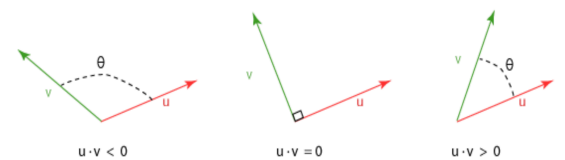
4. unit vector

- vector whose magnitude is one
- the unit vector in the same direction of u : $\hat{u} = \frac{u}{\|u\|}$



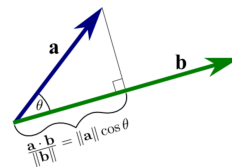
5. dot product geometrically

- $a \cdot b = \|a\| \|b\| \cos \theta$
- if a and b are perpendicular, $a \cdot b = 0$
- if a and b are **opposing** directions, $a \cdot b < 0$
- if a and b are pointing in the same direction, $a \cdot b > 0$



6. orthogonal basis

- basis vectors are orthogonal to each other
- the dot product of any two vectors in the set is zero



7. orthonormal basis

- orthogonal + normalize
- all vectors in the set are **unit** that are orthogonal to each other

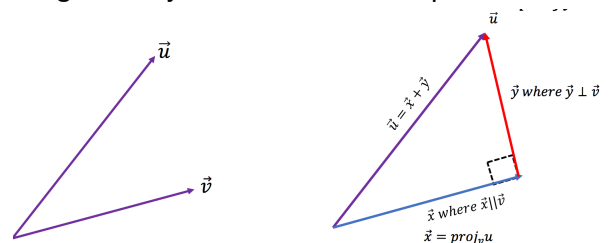
8. **orthogonal matrix** A: A is square matrix with **orthonormal** columns. $A^{-1} = A^T$

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

9. Gram-Schmidt Procedure

- a method to find an orthogonal basis for a subspace given any basis for that subspace
- given basis \vec{u}, \vec{v} , the orthogonal basis \vec{x}, \vec{y}

$$\vec{x} = \text{proj}_{\vec{v}} \vec{u} = \frac{\vec{v} \cdot \vec{u}}{\vec{v} \cdot \vec{v}} \vec{v}; \vec{y} = \vec{u} - \vec{x}$$



1. Eigenvector & Eigenvalue

- **Eigenvector** is a special vector that after linear transformation A, it turns into scalar multiples of itself.
 - they are the ones that stay on their own span during a transformation.
- **Eigenvalue** is the factor by which it is stretched or squashed during the transformation.
- $Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$, x is a **nonzero vector (eigenvector)** and λ is a scalar (eigenvalue)
 - reason: identity matrix multiple any matrix won't change that matrix
- **Characteristic equation:**
 - $\det(A - \lambda I) = 0$
- How to compute eigenvalue?
 - homogeneous system $(A - \lambda I)x = 0$ has **nontrivial/nonzero** solution for particular λ
 - matrix $A - \lambda I$ is not invertible $\Leftrightarrow \det(A - \lambda I) = 0$
 - get λ
- How to compute eigenvector with respect to specific eigenvalue λ ?
 - eigenvector is the solution set of the homogeneous system $(A - \lambda I)x = 0$
 - rref of $A - \lambda I$; write all variables in terms of free variables; parametric form.

2. Basis

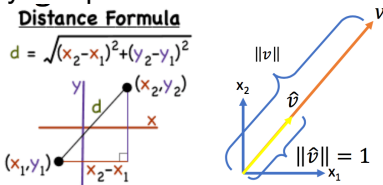
- a basis for a subspace is a linearly independent set whose span is the subspace
- **standard** basis: identity matrix
- a basis for the column space A is formed from the linearly independent columns of A

3. magnitude of a vector:

- the vector's length / the distance from tail to head
- $\|u\| = \sqrt{u \cdot u}$

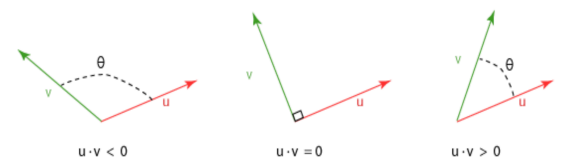
4. unit vector

- vector whose magnitude is one
- the unit vector in the same direction of u : $\hat{u} = \frac{u}{\|u\|}$



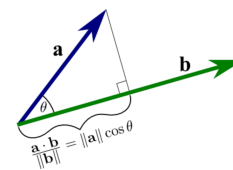
5. dot product geometrically

- $a \cdot b = \|a\| \|b\| \cos \theta$
- if a and b are perpendicular, $a \cdot b = 0$
- if a and b are **opposing** directions, $a \cdot b < 0$
- if a and b are pointing in the same direction, $a \cdot b > 0$



6. orthogonal basis

- basis vectors are orthogonal to each other
- the dot product of any two vectors in the set is zero



7. orthonormal basis

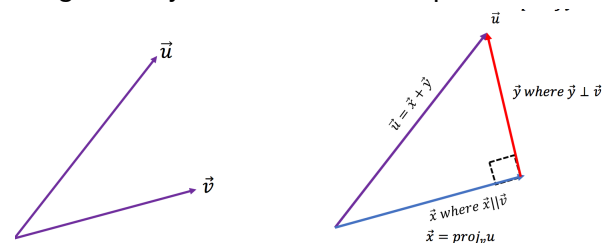
- orthogonal + normalize
- all vectors in the set are **unit** that are orthogonal to each other

8. **orthogonal matrix** A: A is square matrix with **orthonormal** columns. $A^{-1} = A^T$ $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

9. Gram-Schmidt Procedure

- a method to find an orthogonal basis for a subspace given any basis for that subspace
- given basis \vec{u}, \vec{v} , the orthogonal basis \vec{x}, \vec{y}

• $\vec{x} = proj_{\vec{v}} \vec{u} = \frac{\vec{v} \cdot \vec{u}}{\|\vec{v}\|^2} \cdot \vec{v}$; $\vec{y} = \vec{u} - \vec{x}$



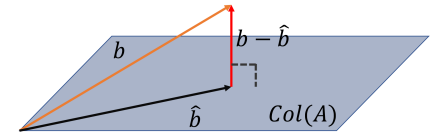
1. Least Square Approximation

- the approximate solution for inconsistent system $Ax=b$
- principle:** b is not in the column space of A , therefore the system has no solution for x .

But it has approximate solution \hat{x} , where the corresponding \hat{b} is closer to b than any other vector in $\text{col}(A)$.

- objective:** minimize the distance between **real** b (not in $\text{col}(A)$) and **approximate** \hat{b} (in $\text{col}(A)$)

- $\min \|b - \hat{b}\|^2$ where $b - \hat{b}$ is the vector going from $\text{col}(A)$ to b

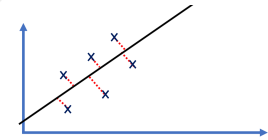


- solution:**

- $\hat{b} = \text{Proj}_{\text{col}(A)} b \Leftrightarrow b - \hat{b}$ is orthogonal to $\text{col}(A)$
- $A^T(b - \hat{b}) = 0 \Leftrightarrow A^T b = A^T \hat{b} \Leftrightarrow A^T A \hat{x} = A^T b$ (matrix-vector multiplication)
- augment matrix $[A^T A | A^T \hat{b}]$, do row reduction, get rref and the solution set for \hat{x}
- \hat{x} is the approximate solution for $Ax=b$

2. Least Square Regression Line

- problem:
 - given a set of points: $\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \dots, \langle x_n, y_n \rangle$
 - find a regression line (a best-fit line) that attempts to fit a set of points to a straight line. Format the regression line $y = C_1 x + C_0$
- objective:** minimize the **mean squared error** between every data point and the line.
 - equals: minimize the average **distance** between a data point and its orthogonal projection on the line
- matrix format: $Ax=b$. no solution.
- find approximate solution using Least Square Approximation (above)



$$\begin{matrix} A & & b \\ \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} & \begin{bmatrix} C_1 \\ C_0 \end{bmatrix} & = & \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \end{matrix}$$

3. Least Square Regression Curve

- order** corresponds to the **highest exponent** of x
- 1st order curve (line): $y = C_1 x + C_0$
- 2nd order curve (parabola): $y = C_2 x^2 + C_1 x + C_0$
- 3rd order curve: $y = C_3 x^3 + C_2 x^2 + C_1 x + C_0$
- n^{th} order curve: $y = C_n x^n + \dots + C_3 x^3 + C_2 x^2 + C_1 x + C_0$
- Theory: a set of n data points can be fit perfectly with a curve of order $n-1$
 - e.g.: three points can be fit perfectly by a parabola.

4. Cross product

- definition: $v \times w = p$, cross product p is a vector
- magnitude:
 - the area of the parallelogram that vector v, w enclose
 - $\det([v, w])$
- direction:
 - p is perpendicular to v and w
 - obeys the right hand rule (the forefinger points to v)

