Revision

COMP408 - Linear Algebra
Dennis Wong

Take away so far

1. Vector
    - Magnitude and Direction
    - Vector addition
    - Vector multiplication

2. System of linear functions
    - Augmented matrix
    - Row echelon form and reduced row echelon form

3. Basis and Dimension
    - Linear independence
    - Coordinate vector
    - Span and subspace

Take away so far

4. Matrix
    - Column space
    - Row space
    - Null space

5. Orthogolity
    - Least square
    - Gram Schmidt

Hopefully give you guys a strong ***foundation*** for what is coming in your future.

Vectors

A ***scalar*** is simply a number, and a ***vector*** is a list of numbers.

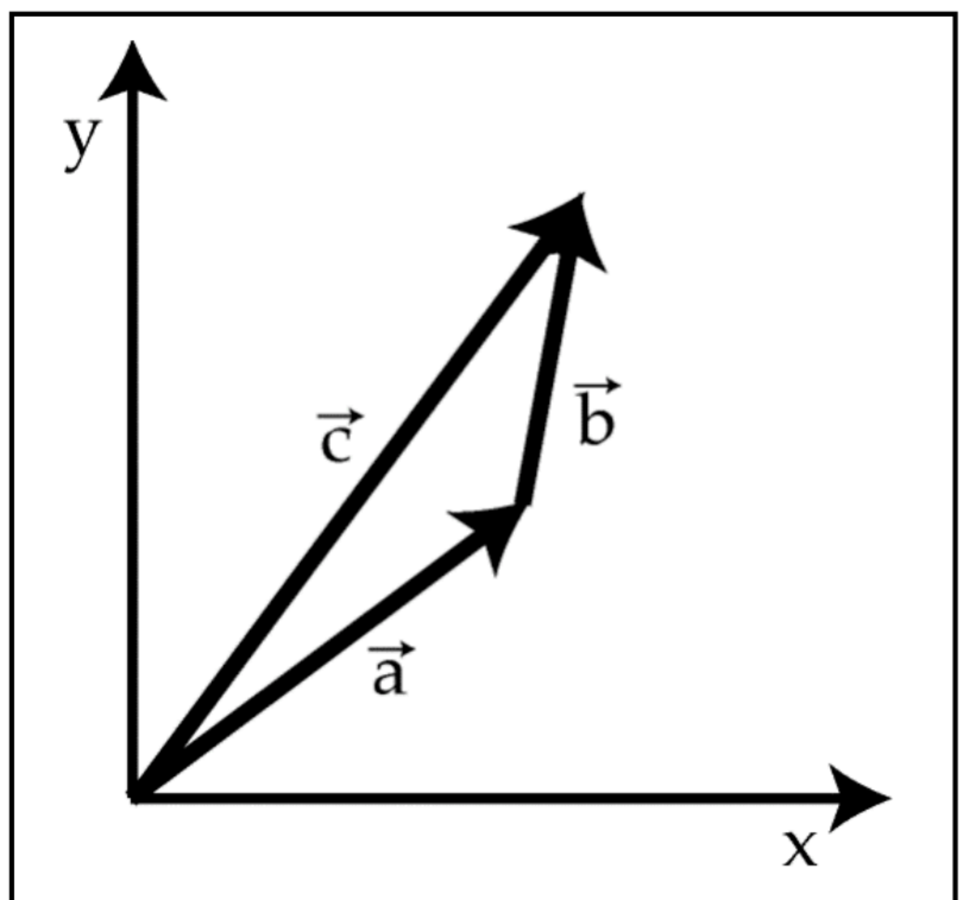
$\vec{a} = [4, 2]$
$\vec{b} = [4, 2, 3, 1]$
$n = 3$

We usually denote a vector with an arrow $\rightarrow$ on top of the symbol.

Vector addition

Vectors can be added and subtracted.

Graphically, we can think of adding two vectors together as placing two line segments end-to-end, maintaining distance and direction.



Vector addition

Numerically, we add vectors component-by-component.

Example: In the morning a ship sailed 4 miles east and 3 miles north, and then in the afternoon it sailed a further 1 mile east and 2 miles north, what was the total displacement for the whole day?

Morning trip: [4, 3]
Afternoon trip: [1, 2]
Total displacement = [4, 3] + [1, 2] = [4+1, 3+2] = [5, 5]

Magnitude and Direction

The ***magnitude*** of a vector is the distance from the endpoint of the vector to the origin, that is, it's ***length***.

The magnitude of a vector $\vec{a}$, denoted by $|\vec{a}|$, can be computed by the Pythagorean theorem.

Example: $\vec{a} = [4, 3]$ and so $|\vec{a}| = \sqrt{(4^2 + 3^2)} = 5$.

A ***unit vector***, denoted by ^ on top, is a vector of magnitude 1. Unit vectors can be used to express the direction of a vector independent of its magnitude.

Example: The unit vector that corresponds to the direction of $\vec{a} = [4, 3]$ is $\hat{a} = [4, 3] / |\vec{a}| = [4/5, 3/5]$.

Augmented matrix

Let's simplify the system of linear equations.

$$x_1 - 2x_2 + 3x_3 = 1$$
$$2x_1 - 3x_2 + 5x_3 = 0$$
$$-x_1 + 4x_2 - x_3 = -1$$

We use a rectangular array of numbers to represent the coefficients of each variable, and the constant term of each linear equation is represented by the numbers on the right.

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 2 & -3 & 5 & 0 \\ -1 & 4 & -1 & -1 \end{array}\right]$$

Row operations

The operations we have been using to reduce a system of equations are called ***row operations***.

We summarize the three types of operation below
1. Interchange two rows
2. Multiply a row by a nonzero number
3. Add a multiple of one row to another row
(4). Add a multiple of one row to a nonzero multiple of another row

The row operations of type 1, 2, and 3 are the ***elementary row operations***. Type 4 is a combination of type 2 and 3.

Two matrices A and B are ***row equivalent*** (written $A \sim B$) if B is obtained from A by applying one or more elementary row operations.

Row operations

Applying row operations on the augmented matrix:

$$
\begin{array}{c@{}l}
\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 2 & -3 & 5 & 0 \\ -1 & 4 & -1 & -1 \end{array}\right] & \begin{array}{l} -2 \quad 1 \\ \\ \\ \end{array}
\end{array}
\sim
\begin{array}{c@{}l}
\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 2 & 2 & 0 \end{array}\right] & \begin{array}{l} \\ -2 \\ \\ \end{array}
\end{array}
$$

$$
\sim
\begin{array}{c@{}l}
\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 4 & 4 \end{array}\right] & \begin{array}{l} \\ \\ 1/4 \end{array}
\end{array}
$$

$$
\sim
\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right]
$$

We can then solve the system by using back substitution as before, which we have x_1 = -4, x_2 = -1 and x_3 = 1.

Row echelon form

A matrix is in ***row echelon form*** if (a) its nonzero rows come before its zero rows, (b) each of its pivot entries is to the right of the pivot entry in the row above (if any).

The first nonzero entry in each nonzero row is called that row's ***pivot*** entry.

$$\begin{bmatrix} 2 & -3 & 0 & 4 \\ 0 & 0 & 7 & 6 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{(REF)} \qquad \begin{bmatrix} 3 & 2 & 7 \\ 1 & 4 & -1 \\ 0 & -8 & 6 \end{bmatrix} \quad \text{(not REF)}$$

Reduced row echelon form

A matrix is in ***reduced row echelon form*** if (a) it is in row echelon form, (b) each entry above (and below) a pivot entry is 0, (c) each pivot entry is 1.

$$\begin{bmatrix} 0 & 0 & 6 & 10 & -1 \\ 3 & 1 & -2 & -5 & -3 \\ 6 & 2 & 0 & -9 & -1 \\ -3 & -1 & 4 & 3 & 8 \end{bmatrix} \overset{\text{Ex. 1.5.1}}{\sim} \begin{bmatrix} 3 & 1 & -2 & -5 & -3 \\ 0 & 0 & 6 & 10 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \\ \\ -10 \quad 5 \\ \\ \end{array}$$

$$\sim \begin{bmatrix} 3 & 1 & -2 & 0 & -8 \\ 0 & 0 & 6 & 0 & 9 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} 3 \\ 1 \\ \\ \\ \end{array}$$

$$\sim \begin{bmatrix} 9 & 3 & 0 & 0 & -15 \\ 0 & 0 & 6 & 0 & 9 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \frac{1}{9} \\ \frac{1}{6} \\ \\ \\ \end{array}$$

$$\sim \begin{bmatrix} 1 & \frac{1}{3} & 0 & 0 & -\frac{5}{3} \\ 0 & 0 & 1 & 0 & \frac{3}{2} \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(RREF)}$$

Reduced row echelon form

Once we have the matrix in reduced row echelon form, it is easy to write down the solution.

There are again three possibilities:
1. Unique solution
2. Infinitely many solutions
3. No solution

Example (unique solution):

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$x_1 = 2$, $x_2 = -3$, and $x_3 = 5$.

Linear independence

A family of vectors is ***linearly independent*** if no one of the vectors can be created by any linear combination of the other vectors in the family.

In other words, if two vectors point in different directions, they are said to be linearly independent.

If two vectors point in the same direction, then we can multiply one of the vector with a scalar to get the other vector, and the two vectors are said to be ***linearly dependent***.

Linear independence

We say that vectors $x_1, x_2, \ldots, x_s$ in R^n are ***linearly dependent*** if there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_s$ not all zero such that

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_s x_s = 0.$$

We say that the vectors are ***linearly independent*** if they are not linearly dependent, that is, if

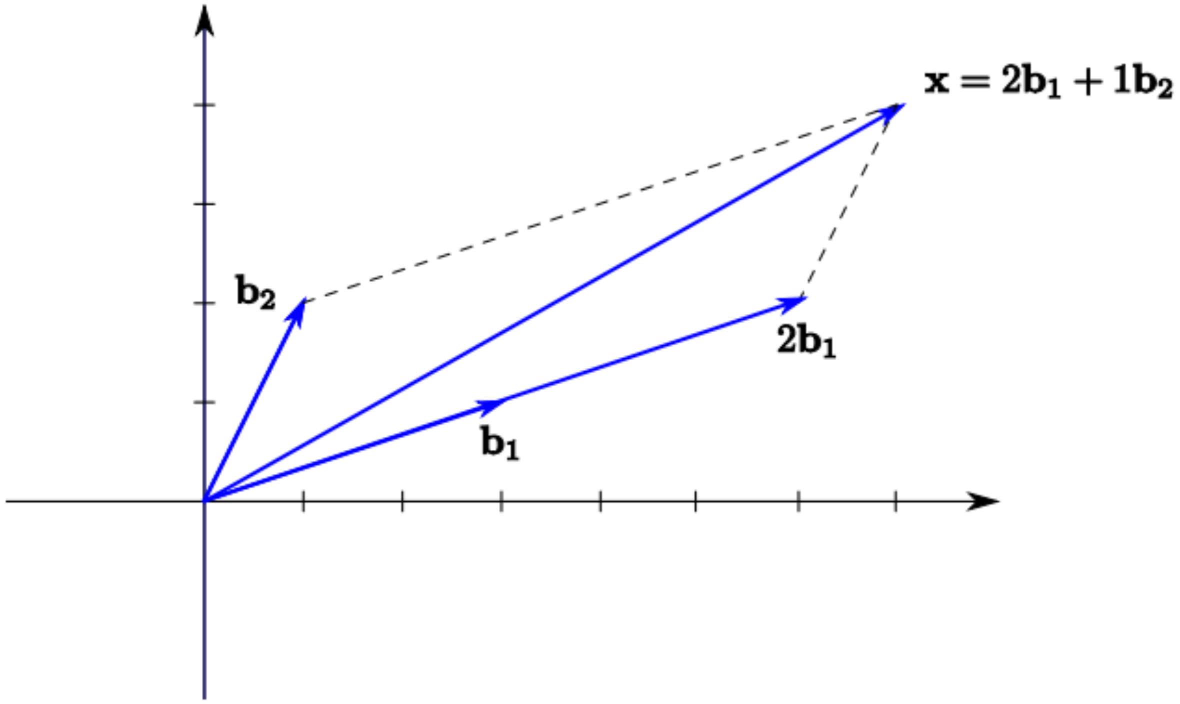
$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_s x_s = 0 \text{ implies } \alpha_i = 0 \text{ for all } i.$$

The zero vector can never be on a list of independent vectors because $\alpha\vec{0} = \vec{0}$ for any scalar α.

Basis

Let S be a subspace of R^n and let $b_1, b_2, . . . , b_s$ be vectors in S. The set $\{b_1, b_2, . . . , b_s\}$ is a ***basis*** for S if

1. $\text{Span}\{b_1, b_2, . . . , b_s\} = S$.
2. $b_1, b_2, . . . , b_s$ are linearly independent.



Any vector in S can be written ***uniquely*** as a linear combination of the vectors in basis.

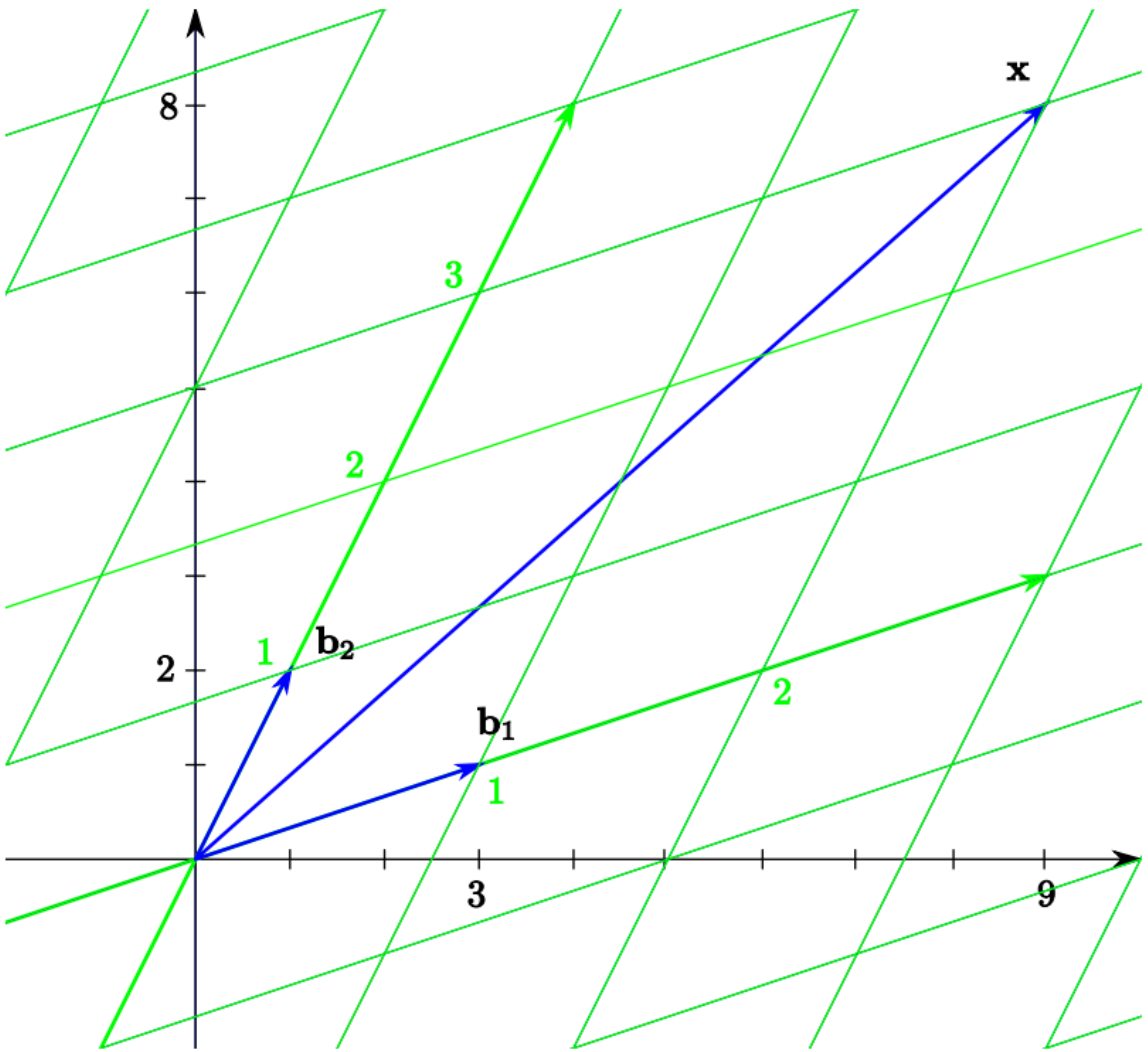
Coordinate vector

Let S be a subspace of R^n, let B = (b_1, b_2, . . . , b_s) be an ordered basis for S, and let x be a vector in S. The ***coordinate vector of*** x ***relative to B*** is

$$[x]_B = [a_1, a_2, \ldots, a_s],$$

where x = $a_1b_1 + a_2b_2 + \ldots + a_sb_s$.

Coordinate vector



Dimension

Let S be a subspace of R^n. If S has a basis consisting of s vectors, we say that S has ***dimension*** s and we write dim $S = s$.

Example: Find the dimension of S = Span{x_1, x_2} where x_1 = [1, 1, 0] and x_2 = [0, 1, 1].

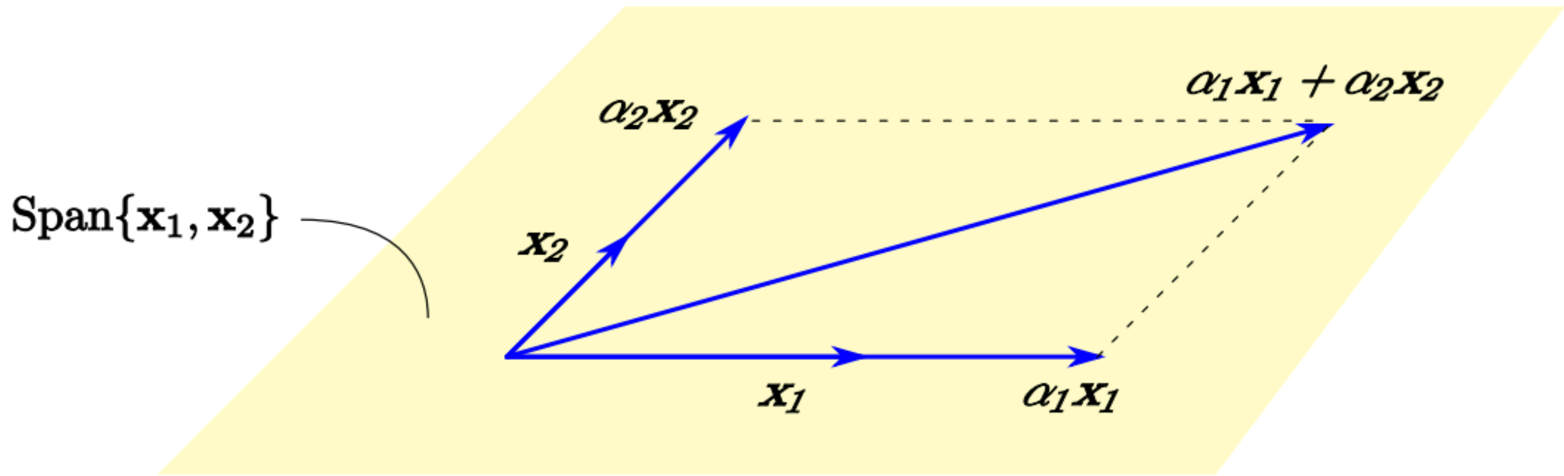
We can prove that x_1 and x_2 are linearly independent and also thus x_1 and x_2 are basis of S. Therefore the dimension of S is 2.

Span

Let $\{x_1, x_2, \ldots, x_s\}$ be a set of vectors in $\mathbf{R}^n$. The ***span*** of $\{x_1, x_2, \ldots, x_s\}$, denoted by $\mathrm{Span}\{x_1, x_2, \ldots, x_s\}$, is the set of all linear combinations of $x_1, x_2, \ldots, x_s$:

$$\mathrm{Span}\{x_1, x_2, \ldots, x_s\} = \{\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_s x_s \mid \alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbf{R}\}.$$

If x_1 and x_2 are not parallel, then one can show that $\mathrm{Span}\{x_1, x_2\}$ is the ***plane*** determined by x_1 and x_2.



Span

We can use system of linear equations to determine if a vector is in a span or not.

Example: Determine whether $[2, -5, 8]^T$ is in Span$\{x_1, x_2\}$.

$$\begin{bmatrix} 2 \\ -5 \\ 8 \end{bmatrix} = \alpha_1 \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 + 4\alpha_2 \\ -\alpha_1 + 2\alpha_2 \\ 3\alpha_1 + \alpha_2 \end{bmatrix} .$$

Equating components leads to the following augmented matrix:

$$\left[\begin{array}{cc|c} 2 & 4 & 2 \\ -1 & 2 & -5 \\ 3 & 1 & 8 \end{array}\right]\begin{array}{cc} 1 & -3 \\ 2 & \\ & 2 \end{array} \sim \left[\begin{array}{cc|c} 2 & 4 & 2 \\ 0 & 8 & -8 \\ 0 & -10 & 10 \end{array}\right]\begin{array}{l} \\ \frac{1}{8} \\ \frac{1}{10} \end{array}$$

$$\sim \left[\begin{array}{cc|c} 2 & 4 & 2 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{array}\right]\begin{array}{l} \\ 1 \\ \end{array}$$

$$\sim \left[\begin{array}{cc|c} 2 & 4 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right] .$$

Column space and row space

The ***column space (range space)***, col(A), of A is the subspace of R^m spanned by the columns of A.

The ***row space***, row(A), of A is the subspace of R^n spanned by the rows of A.

$$\begin{bmatrix} 1 & 8 & 13 & 12 \\ 14 & 11 & 2 & 7 \\ 4 & 5 & 16 & 9 \\ 15 & 10 & 3 & 6 \end{bmatrix} \qquad \begin{bmatrix} 1 & 8 & 13 & 12 \\ 14 & 11 & 2 & 7 \\ 4 & 5 & 16 & 9 \\ 15 & 10 & 3 & 6 \end{bmatrix}$$

If A and B are matrices with $A \sim B$, then row(A) = row(B).

Bases of column space and row space

If R is a row-echelon matrix, then

1. The nonzero rows of R are a basis of row R.
2. The columns of R containing leading ones are a basis of col R.

Example: Consider the following matrix

$$\begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 7 & 3 & 9 \\ 1 & 5 & 3 & 1 \\ 1 & 2 & 0 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & -3 \\ 0 & -1 & -1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The basis of the column space are $(1, 2, 1, 1)^T$, $(3, 7, 5, 2)^T$ and $(4, 9, 1, 8)^T$.

Null space and image space

The null space of A, denoted null(A), is defined by null(A) = $\{x \in R^n \mid Ax = 0\}$.

The image space of A, denoted $im(A)$, are defined by $im(A) = \{Ax \mid x \in R^n\}$

In other words, $null(A)$ consists of all solutions x in R^n of the homogeneous system $Ax = 0$, and $im(A)$ is the set of all vectors y in R^m such that $Ax = y$ has a solution x.

Dot Product

A ***dot product*** (or ***inner product***) is the numerical product of the lengths of two vectors, multiplied by the cosine of the angle between them, that is $\vec{a} \cdot \vec{b} = |\vec{a}|\ |\vec{b}|\ cos\ \theta$, where θ represents the angle between the two vectors.

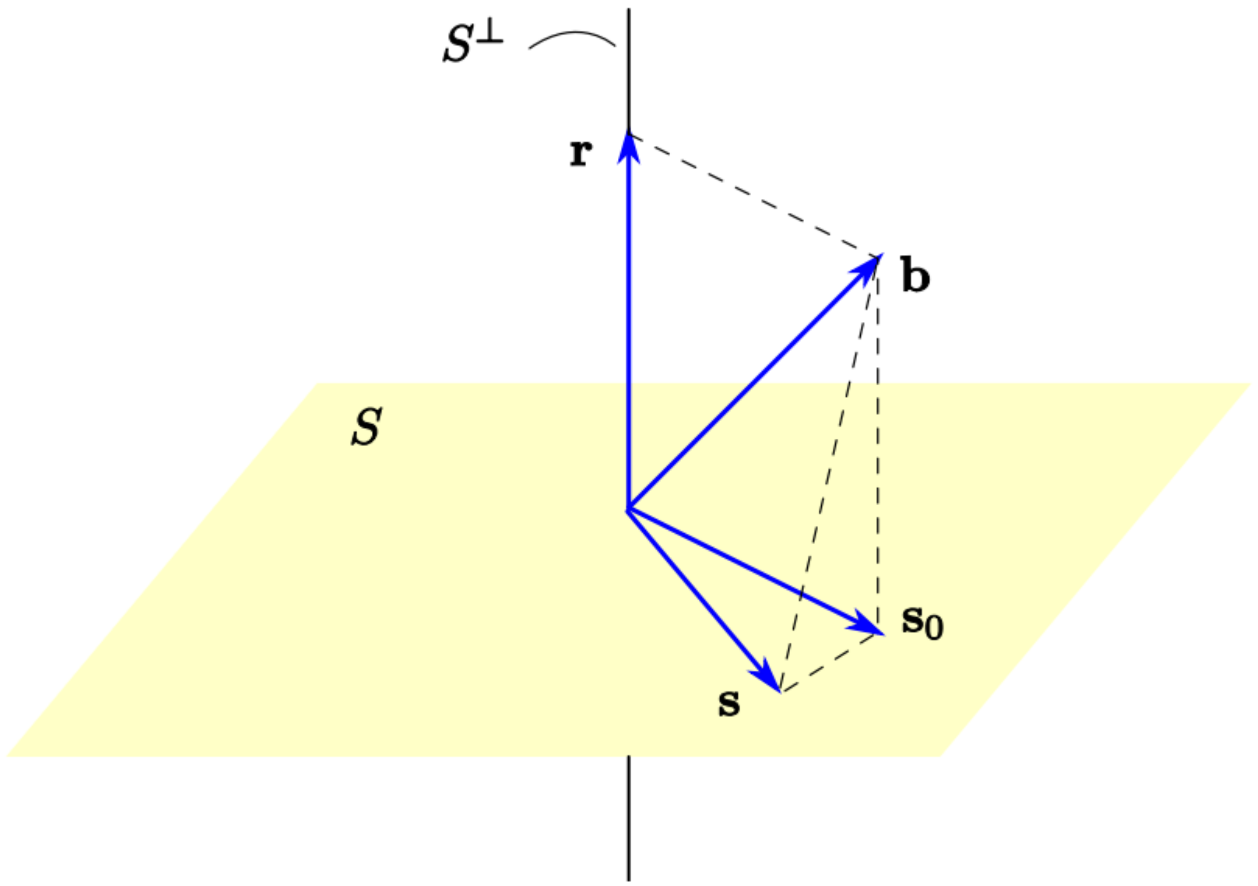
A simply way to calculate a dot product is by multiplying the components of each vector separately and then adding these products together.

Example: $\vec{a} = [4, 3], \vec{b} = [1, 2]$
$\vec{a} \cdot \vec{b} = (4 \text{ x } 3) + (3 \text{ x } 2) = 11$

Least squares

Let S be a subspace of V, let b be a vector in V and assume that $b = s_0 + r$ with $s_0 \in$ S and r $\in S^\perp$.

For every $s \in$ S we have dist(b, s_0) $\leq$ dist(b, s).



Least squares

Suppose that the matrix equation $Ax = b$ has no solution. In terms of distance, this means that dist(b, Ax) is never zero, no matter what x is.

Instead of leaving the equation unsolved, it is sometimes useful to find an x_0 that is as close to being a solution as possible.

That is, for which the distance from b to Ax_0 is less than or equal to the distance from b to Ax for every other x. This is called a ***least squares solution***.

Least squares

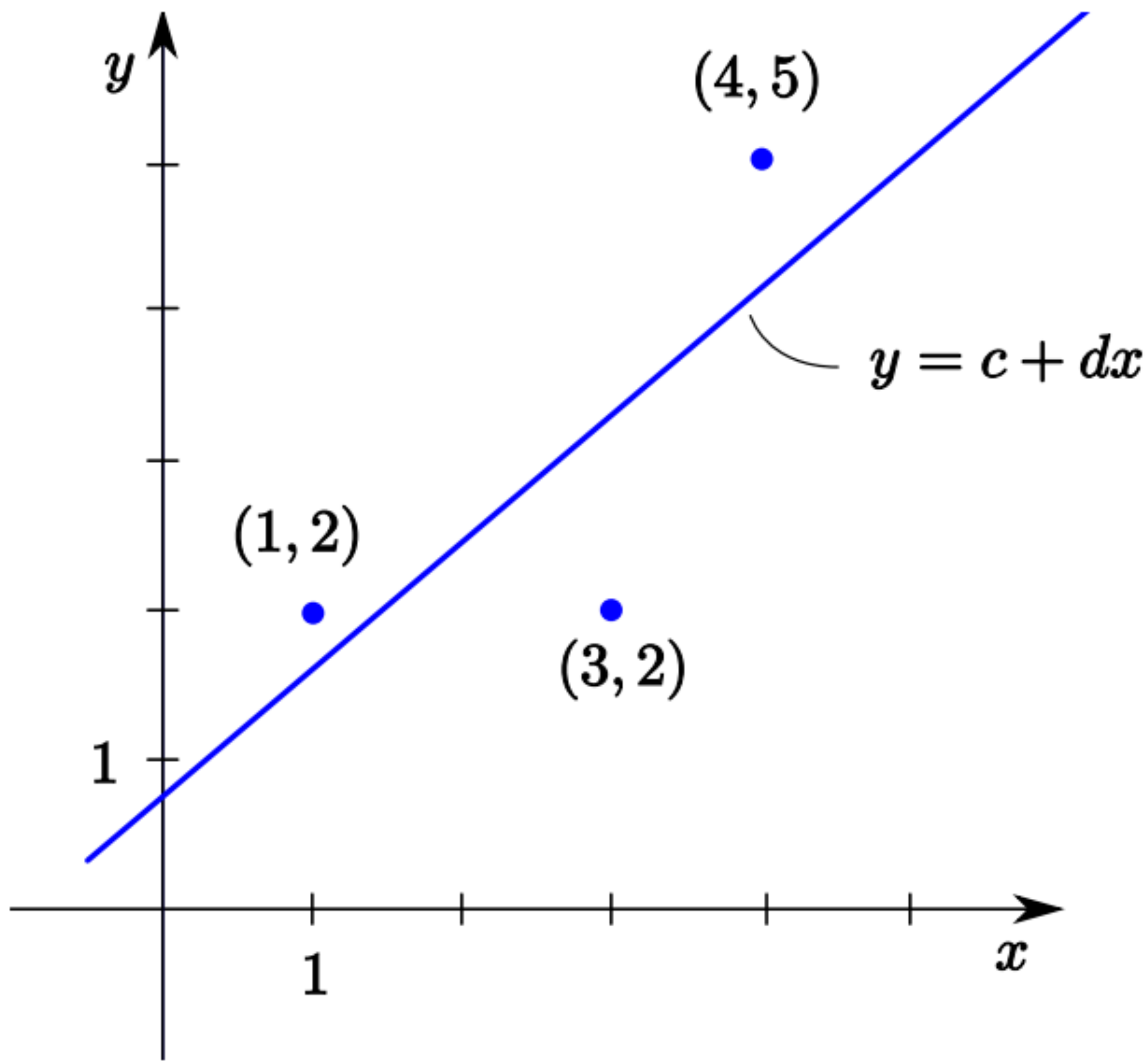
Let A be an $m \times n$ matrix and let b be a vector in R^n. If $x = x_0$ is a solution to

$$A^T A x = A^T b,$$

then, for every $x \in \mathrm{R}^n$, $\mathrm{dist}(b, Ax_0) \leq \mathrm{dist}(b, Ax)$.

Such an x_0 is called a least squares solution to the equation $Ax = b$.

Least squares



Least squares

Example: Use a least squares solution to find a line that best fits the data points (1, 2), (3, 2), and (4, 5).

If we write the desired line as c+dx = y, then ideally the line would go through all three points giving the system

$$\begin{aligned} c + d &= 2 \\ c + 3d &= 2 \\ c + 4d &= 5 \end{aligned}$$

which can be written as the matrix Ax = b with

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}.$$

Least squares

Applying least square solution by solving the equation $A^{\mathsf{T}}Ax = A^{\mathsf{T}}b$, we have the following:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 8 & 26 \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 28 \end{bmatrix},$$

Least squares

Solving the matrix solution gives us the following:

$$
\begin{array}{c}\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 8 & 26 & 28 \end{array}\right]\end{array}\begin{array}{c} -8 \\ 3 \end{array}
\sim
\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 0 & 14 & 12 \end{array}\right]\begin{array}{c} \\ \frac{1}{2} \end{array}
$$

$$
\sim
\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 0 & 7 & 6 \end{array}\right]\begin{array}{c} 7 \\ -8 \end{array}
$$

$$
\sim
\left[\begin{array}{cc|c} 21 & 0 & 15 \\ 0 & 7 & 6 \end{array}\right]\begin{array}{c} \frac{1}{21} \\ \frac{1}{7} \end{array}
$$

$$
\sim
\left[\begin{array}{cc|c} 1 & 0 & \frac{5}{7} \\ 0 & 1 & \frac{6}{7} \end{array}\right]
$$

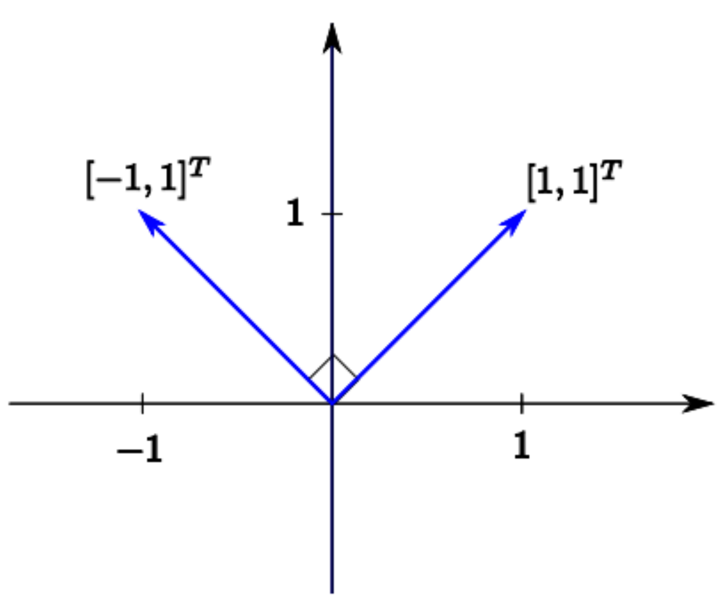
That is, c = 5/7 and d = 6/7 and the best fitting line is y = 5/7 + 6/7x.

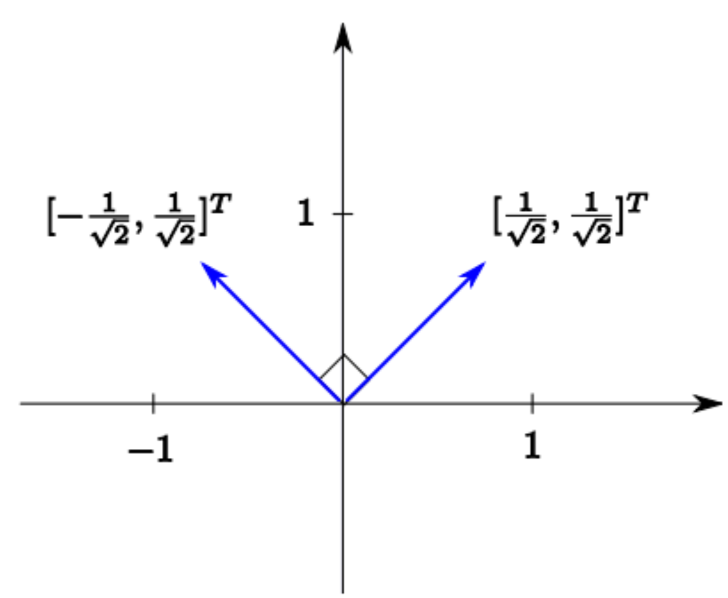
Orthogonal set

Let $\{b_1, b_2, \ldots, b_s\}$ be a set of vectors in the inner product space V. The set is ***orthogonal*** if $\{b_i, b_j\} = 0$ for all $i \neq j$ (the vectors are ***pairwise orthogonal***).

The set is ***orthonormal*** if it is orthogonal and each vector is a unit vector.

Any orthogonal set of nonzero vectors can be changed into an orthonormal set by dividing each vector by its norm.





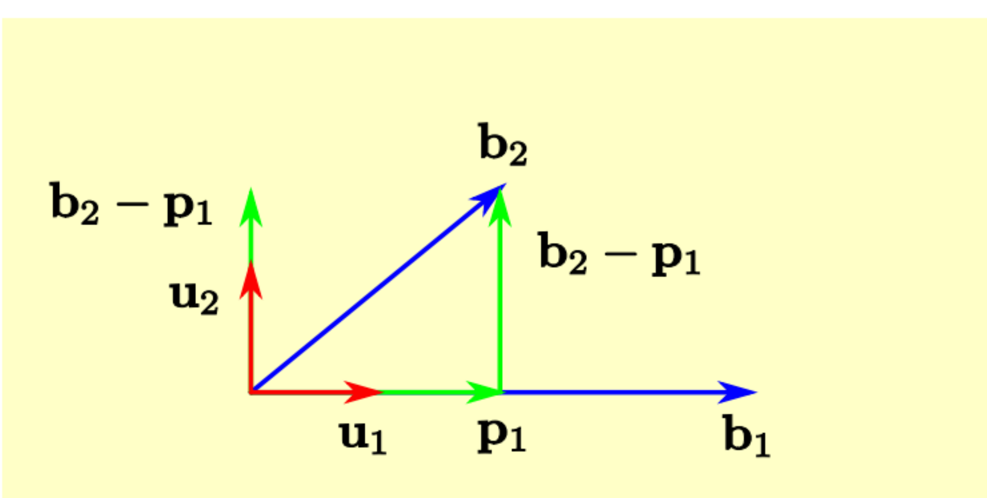
Gram-Schmidt

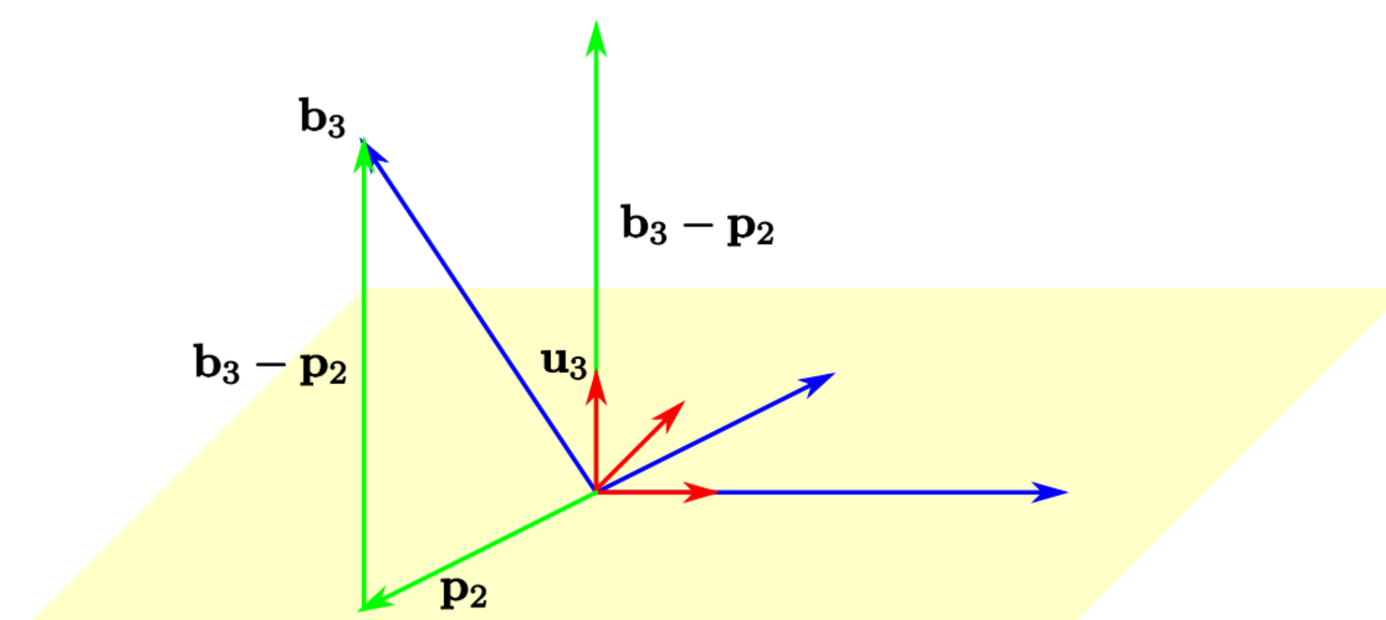
Let $\{b_1, b_2, \ldots, b_s\}$ be a basis for the inner product space V. Define vectors $u_1, u_2, \ldots, u_2$ recursively by

$$\mathbf{u}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$$

$$\mathbf{u}_k = \frac{\mathbf{b}_k - \mathbf{p}_{k-1}}{\|\mathbf{b}_k - \mathbf{p}_{k-1}\|}, \quad \text{where } \mathbf{p}_{k-1} = \sum_{i=1}^{k-1} \langle \mathbf{b}_k, \mathbf{u}_i \rangle \mathbf{u}_i \quad (k > 1)$$

Then $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis for V. Moreover, $\mathrm{Span}\{u_1, u_2, \ldots, u_k\} = \mathrm{Span}\{b_1, b_2, \ldots, b_k\}$ for each k.





Gram-Schmidt

Example: Let $b_1 = [1, 2, 2, 4]^T$, $b_2 = [-2, 0, -4, 0]^T$, and $b_3 = [-1, 1, 2, 0]^T$, and let S be the span of these vectors. Apply the Gram-Schmidt process to $\{b_1, b_2, b_3\}$ to obtain an orthonormal basis $\{u_1, u_2, u_3\}$ for S.

Solution: First we compute u_1 and p_1:

$$u_1 = b_1 / \| b_1 \| = [1, 2, 2, 4]^T / \| [1, 2, 2, 4]^T \| = 1/5[1, 2, 2, 4]^T$$

$$\begin{aligned} p_1 = \langle b_2, u_1 \rangle u_1 &= \langle [-2, 0, -4, 0]^T, 1/5[1, 2, 2, 4]^T \rangle u_1 \\ &= -2/5[1, 2, 2, 4]^T \end{aligned}$$

Gram-Schmidt

Solution (cont): Then, we compute $b_2 - p_1$ and u_2:

$b_2 - p_1 = [-2, 0, -4, 0]^T + 2/5[1, 2, 2, 4]^T = 4/5[-2, 1, -4, 2]^T$

$u_2 = (b_2 - p_1) \|b_2 - p_1\| = 4/5[-2, 1, -4, 2]^T / \| 4/5[-2, 1, -4, 2]^T \|$
$= 1/5[-2, 1, -4, 2]^T$

Finally we compute p_2, $b_3 - p_2$, and u_3:

$p_2 = <b_3, u_1> u_1 + <b_3, u_2> u_2$
$= <[-1, 1, 2, 0]^T, 1/5[1, 2, 2, 4]^T> u_1 + <[-1, 1, 2, 0]^T, 1/5[-2, 1, -4, 2]^T> u_2$
$= 1/5[3, 1, 6, 2]^T$

$b_3 - p_2 = [-1, 1, 2, 0]^T - 1/5[3, 1, 6, 2]^T = 2/5[-4, 2, 2, -2]^T$

$u_3 = (b_3 - p_2) \|b_3 - p_2\| = 2/5[-4, 2, 2, -1]^T / \| 2/5[-4, 2, 2, -1]^T \|$
$= 1/5[-4, 2, 2, -1]^T$