



Lecture 2

Math Foundations Team



BITS Pilani

Pilani | Dubai | Goa | Hyderabad

- ▶ We already have some notion about the structure of a vector space, i.e adding two vectors returns a vector, multiplying a vector by a scalar also returns a vector.
- ▶ To formalize these notions, we need the concept of a *Group*
- ▶ *A Group is a set G and an operation $\otimes : G \times G \rightarrow G$ defined on G . Then (G, \otimes) is called a group if the following properties hold:*
 - ▶ *Closure of G under \otimes : $\forall x, y \in G, x \otimes y \in G$*
 - ▶ *Associativity: $\forall x, y, z \in G, (x \otimes y) \otimes z = x \otimes (y \otimes z)$*
 - ▶ *Neutral or Identity element: $\exists e \in G, \forall x \in G, x \otimes e = x$*
 - ▶ *Inverse element: $\forall x \in G, \exists y \in G, x \otimes y = y \otimes x = e$*
- ▶ *If in addition to all the above properties we have $\forall x, y \in G, x \otimes y = y \otimes x$ then (G, \otimes) is an Abelian group*



- ▶ $(\mathbb{Z}, +)$ is an Abelian group. We can see that all the aforementioned properties hold.
- ▶ What about $(\mathbb{N}_0, +)$? This is not a group since there is no inverse element for an arbitrary element in it.
- ▶ (\mathbb{Z}, \cdot) where \mathbb{Z} is the set of integers, and \cdot is product. It has the identity element but there exist elements in it that don't have an inverse.



- ▶ We will now consider sets that are just like groups in terms of its properties with respect to an inner operation $+$, and also have an outer operation called $.$ which denotes the multiplication of a vector $x \in G$ by a scalar $\lambda \in R$.
- ▶ The inner operation can be viewed as a form of addition while the outer operation can be viewed as a form of scaling.
- ▶ A real-valued vector space is a set $V = (\mathcal{V}, +, .)$ with operations $+, .$ such that $+: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ and $.: R \times \mathcal{V} \rightarrow \mathcal{V}$ where $(\mathcal{V}, +)$ is an Abelian group, and the following properties hold:
 - ▶ Distributivity
 - ▶ associativity with respect to the outer operation.
 - ▶ there exists a neutral or identity element with respect to the outer operation.



Let us look at the properties of a Vector Space more carefully:

- ▶ Distributivity: $\forall \lambda \in R, \mathbf{x}, \mathbf{y} \in \mathcal{V}, \lambda.(\mathbf{x} + \mathbf{y}) = \lambda.\mathbf{x} + \lambda.\mathbf{y}$ and $\forall \lambda, \psi \in R, (\lambda + \psi).\mathbf{x} = \lambda.\mathbf{x} + \psi.\mathbf{x}$
- ▶ Associativity with respect to the outer operation: $\forall \lambda, \psi \in R, \mathbf{x} \in \mathcal{V}, \lambda.(\psi.\mathbf{x}) = (\lambda\psi).\mathbf{x}$
- ▶ Neutral element with respect to the outer operation: $\forall \mathbf{x} \in \mathcal{V}, 1.\mathbf{x} = \mathbf{x}$
- ▶ The elements $\mathbf{x} \in \mathcal{V}$ are called vectors.
- ▶ The neutral element with respect to $(\mathcal{V}, +)$ is the zero vector $[0, 0, \dots, 0]^T$ and the inner operation is called vector addition.



Consider some examples of vector spaces, for example $\mathcal{V} = \mathbb{R}^n$

- ▶ We can define addition:

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

- ▶ Multiplication by scalars:

$$\lambda \mathbf{x} = \lambda (x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n), \forall \lambda \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n.$$

What we call a vector need not be the standard column vector that we are accustomed to treating as a vector. We can think of $m \times n$ matrices as vectors and create a vector space out of them. Thus $\mathcal{V} = \mathbb{R}^{m \times n}$ with addition and multiplication defined as below:

- Addition $\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} \dots a_{1n} + b_{1n} \\ \vdots \\ a_{n1} + b_{n1} \dots a_{nn} + b_{nn} \end{bmatrix}$ is defined element-wise for two matrices $A, B \in \mathcal{V}$.

- Multiplication by scalars: $\lambda \mathbf{A} = \begin{bmatrix} \lambda a_{11} \dots \lambda a_{1n} \\ \vdots \\ \lambda a_{n1} \dots \lambda a_{nn} \end{bmatrix}$



- ▶ Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and let $\mathcal{U} \subseteq \mathcal{V}, \mathcal{U} \neq \Phi$. then $U = (\mathcal{U}, +, \cdot)$ is called a vector subspace of V if U is a vector space with the vector space operations $+$ and \cdot restricted to $\mathcal{U} \times \mathcal{U}$ and $R \times \mathcal{U}$.
- ▶ We use the notation $U \subseteq V$ to denote that U is a vector subspace of V .
- ▶ If $U \subseteq V$ and V is a vector space, then U naturally inherits many properties directly from V because they hold for all $\mathbf{x} \in V$, and in particular for all $\mathbf{x} \in U \subseteq V$. These properties include the Abelian group property, associativity, distributivity and the neutral element.

How do we show if $(\mathcal{U}, +, \cdot)$ is subspace of V ?

We need to show that

- ▶ $\mathcal{U} \neq \emptyset$
- ▶ Closure of \mathcal{U} with respect to the outer and inner operations, i.e $\forall \lambda \in R, \forall \mathbf{x} \in \mathcal{U}, \lambda \mathbf{x} \in \mathcal{U}$ and $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U}, \mathbf{x} + \mathbf{y} \in \mathcal{U}$.

- ▶ Let $\mathcal{V} = \mathbb{R}^2$ and \mathcal{U} be the y -axis. Is $U = (\mathcal{U}, +, \cdot)$ a subspace of $V = (\mathcal{V}, +, \cdot)$? Answer: Yes, because the addition of any two vectors on the y -axis remains on the y -axis so closure with respect to the addition operation is satisfied. Also any vector on the y -axis when scaled by any real number (including 0) will yield a vector on y -axis.
- ▶ What about when we shift the y -axis one unit to the right, i.e $\mathcal{U} = \{x = 1\}$? Answer: We no longer have a vector subspace since scaling with respect to the outer operation does not have the closure property.



- ▶ What about the subset of \mathbb{R}^2 that represents a square around the origin, i.e. $-1 \leq x \leq 1, -1 \leq y \leq 1$? This is again not a subspace.
- ▶ This subspace is of particular interest to us called the nullspace of a matrix, i.e. the set of solutions to a linear system of equations $A\mathbf{x} = 0$ and denoted by $N(\mathbf{A})$. Why is this a vector subspace?

Linear combination and linear independence



- ▶ We know that we can stay in a vector space by adding vectors belonging to the space, and scaling them?
- ▶ We are now interested in a different question → can we come up with a set of vectors such that every vector in the vector space can be represented as a sum of these vectors with scaling as necessary?
- ▶ The answer is yes. A set of vectors capable of representing all vectors in a vector space is called a basis.
- ▶ To explore the question of finding a basis, we need to learn the concepts of linear combination and linear independence.



- ▶ Consider a vector space V and a set of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in V$. Then every $\mathbf{v} \in V$ that is of the form $\mathbf{v} = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_k \mathbf{x}_k$ is a linear combination of the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$.
- ▶ Note that the $\mathbf{0}$ can be written trivially as a linear combination of the given k vectors \mathbf{x}_i s. We are however interested in non-trivial linear combinations of vectors to get the $\mathbf{0}$ -vector.

Linear (in)dependence



- ▶ Consider a vector space V and k vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in V$. Then if there is a non-trivial linear combination of the given vectors such that $0 = \sum_{i=1}^k \lambda_i \mathbf{x}_i$ where at least one of the λ_i s is non-zero, then the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are linearly dependent.
- ▶ Put another way, if the only way we can combine vectors to get the **0**-vector is by letting all the λ_i be zero, then the vectors are linearly independent.
- ▶ What does the concept of linear independence capture? If the vectors are linearly dependent we can write one of the vectors in terms of the others, so that vector is redundant. On the other hand, when the vectors are linearly independent, each vector brings something to the table which the other vectors collectively cannot replace.



- ▶ k vectors in a vector space are either linearly independent or linearly dependent. There is no third option.
- ▶ If at least one of the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ is the 0-vector, then the vectors are linearly dependent. Why? Apply the definition to see that this is the case. We can choose a non-zero λ for the 0-vector and zero λ s for all the other vectors to get the linear combination to be zero.
- ▶ If all the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are non-zero, then the vectors are linearly dependent if and only if one of them is a linear combination of the others.

Using Gaussian elimination to check for linear independence



- ▶ A practical way of checking whether a bunch of vectors are linearly independent is to fill out the columns of a matrix with the given vectors and then perform Gaussian elimination to get the row-echelon form
- ▶ The pivot columns indicate all the vectors that are linearly independent, and they are all on the left.
- ▶ The non-pivot columns can be expressed as a linear combination of the pivot columns

Example



- ▶ Let us start with the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \end{bmatrix}$$

- ▶ Gaussian elimination of this matrix will give rise to the following row-echelon form

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

- ▶ The pivot columns are the first and the third column and we see that the second column which is a non-pivot column can be expressed as a linear combination of the pivot columns to its left, which is just twice the first column.

Another example



Consider the vectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}$

To check for linear independence we set up the equation with λ s as follows: $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \lambda_3 \mathbf{x}_3 = \mathbf{0}$.

As before we put the given vectors into a matrix to get:

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix}$$



We now perform Gaussian elimination on this matrix to identify pivot and non-pivot columns. If all columns are pivot columns then the given three vectors are linearly independent. After Gaussian elimination we get

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

Note that all columns are pivot columns, so we conclude that the original three vectors are linearly independent. The only way to combine those vectors and get the 0-vector is to take 0s as the λ s.



- ▶ Let us say that we have a bunch of k independent vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k$
- ▶ We now have m vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$, each of which is some linear combination of the vectors \mathbf{b}_i .
- ▶ The goal is to find under what conditions the vectors \mathbf{x}_i are linearly independent.
- ▶ Can we express the criterion of linear independence of the \mathbf{x}_i s in terms of the way in which we combine the \mathbf{b}_i s to get the \mathbf{x}_i s?



We have the following equations:

$$\mathbf{x}_1 = \sum_{i=1}^{i=k} \lambda_{i1} \mathbf{b}_i;$$

$$\mathbf{x}_2 = \sum_{i=1}^{i=k} \lambda_{i2} \mathbf{b}_i;$$

$$\vdots$$

$$\mathbf{x}_m = \sum_{i=1}^{i=k} \lambda_{im} \mathbf{b}_i;$$



- ▶ We can define a matrix $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k]$ whose columns are the original linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k$ and then we see that $\mathbf{x}_j = \mathbf{B}\boldsymbol{\lambda}_j$ where $\boldsymbol{\lambda}_j$ is a vector of coefficients.
- ▶ Now to test whether $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are linearly independent, we set up the equation $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}$ and check whether the ψ_j s have to be zero. This means

$$\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \boldsymbol{\lambda}_j = \mathbf{B} \sum_{j=1}^m \psi_j \boldsymbol{\lambda}_j$$



Now $\sum_{j=1}^{j=m} \psi_j \lambda_j = \mathbf{v}$, a vector so $\mathbf{B}\mathbf{v}$ is a linear combination of the column vectors of \mathbf{B} , and the only way a linear combination of the columns of \mathbf{B} can be equal to zero if all the combining coefficients are zero. This means that $\mathbf{v} = \mathbf{0} = \sum_{j=1}^{j=m} \psi_j \lambda_j$.

The fact that $\sum_{j=0}^{j=m} \psi_j \mathbf{x}_j = 0$ implies that $\sum_{j=1}^{j=m} \psi_j \lambda_j = 0$ means that $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are linearly independent if and only if $\lambda_1, \lambda_2, \dots, \lambda_m$ are linearly independent.

Example



Consider the following set of equations:

$$x_1 = \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4$$

$$x_2 = -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4$$

$$x_3 = 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4$$

$$x_4 = 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4$$

We would like to find if the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are linearly independent. Following the theory established in the last slides we check if the vectors corresponding to the λ_j s are linearly independent.

We therefore need to check whether the vectors

$$\begin{bmatrix} 1 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ -2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 17 \\ -10 \\ 11 \\ 1 \end{bmatrix} \quad (1)$$

are linearly independent.

Example



We form a matrix consisting of all these column vectors:

$$A = \begin{bmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & 4 & -3 & 1 \end{bmatrix}$$

and perform Gaussian elimination to get the reduced row-echelon form.

We get the following matrix:

$$B = \begin{bmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Clearly there is a non-trivial way of combining the columns of the above matrix to get the **0**-vector, i.e take $7 \times$ the first column, add it to $15 \times$ the second column, add the result to $18 \times$ the third column and add the result to the fourth column to get **0**. Thus the original set of vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are not linearly independent.



- ▶ We had a set of k vectors \mathbf{b}_i which we combined linearly to get m vectors \mathbf{x}_i .
- ▶ Determining whether the m vectors \mathbf{x}_i were linearly independent boiled down to checking if the column vectors of a $k \times m$ matrix were linearly independent.
- ▶ We performed Gaussian elimination on this matrix to get the reduced row-echelon form.



- ▶ Here is an interesting insight: if $m > k$, we have more columns than rows and there can only be as many pivots as there are non-zero rows, so the number of pivots and therefore the number of pivot columns is less than k .
- ▶ We are therefore **guaranteed** to have non-pivot columns in this case, so that we can express the non-pivot columns in terms of the pivot columns to their left and get a non-trivial linear combination of the columns to get the **0** vector. This would make the m vectors \mathbf{x}_i **linearly dependent**.



- ▶ Do there exist a set of vectors in a vector space which "span" the entire space?
- ▶ What we mean here is that any vector $v \in V$ can be generated as a linear combination of the vectors in question.
- ▶ Is such a set of vectors unique to the vector space?
- ▶ Are all sets capable of generating all vectors in a vector space of the same size?



- ▶ Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and a set of vectors $\mathcal{A} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\} \subseteq V$.
- ▶ If every vector $v \in V$ can be expressed as a linear combination of the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$, then \mathcal{A} is called a **generating set** of \mathcal{V} .
- ▶ The set of all linear combinations of the vectors in \mathcal{A} is known as the **span** of \mathcal{A} .
- ▶ If \mathcal{A} spans the vector space V we write $V = \text{span}[\mathcal{A}]$.
- ▶ If the elements of \mathcal{A} are linearly independent and $\text{span}[\mathcal{A}] = V$, then \mathcal{A} is called the **basis** of V .



- ▶ Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and $\mathcal{A} \subseteq \mathcal{V}$.
- ▶ A generating set \mathcal{A} of V is called minimal if there is no smaller set $\tilde{\mathcal{A}}$ such that $\tilde{\mathcal{A}} \subset \mathcal{A} \subseteq \mathcal{V}$ that spans V .
- ▶ Every linearly independent generating set of V is minimal and is called a basis of V .

Basis is not unique



- ▶ Consider the space \mathbb{R}^3 . The canonical basis is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
- ▶ Another basis for \mathbb{R}^3 is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- ▶ Yet another basis for \mathbb{R}^3 is $\begin{bmatrix} 0.5 \\ 0.8 \\ 0.4 \end{bmatrix}$, $\begin{bmatrix} 1.8 \\ 0.3 \\ 0.3 \end{bmatrix}$, $\begin{bmatrix} -2.2 \\ -3.3 \\ 1.5 \end{bmatrix}$

- ▶ Is any linear independent set a basis?
- ▶ No, we can have a linearly independent set that has too few vectors to become a basis
- ▶ As an example, consider the following set of vectors in \mathbb{R}^4

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 4 \end{bmatrix}$$

- ▶ Why is the above not a basis?



- ▶ Vector spaces can be finite or infinite-dimensional.
- ▶ We consider only finite-dimensional vector spaces.
- ▶ In a finite-dimensional vector space V , the number of vectors in the basis is known as $\dim(V)$.
- ▶ If $U \subseteq V$, then $\dim(U) \leq \dim(V)$.
- ▶ $\dim(U) = \dim(V)$ if and only if $U = V$.



- ▶ Is the size of the basis merely the number of components in the vector?
- ▶ So far, that has been the case.
- ▶ But consider the vector space $\text{span}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$ - the vector has two components, but the basis is just one vector. All vectors in this vector space are multiples of the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.



There are three steps to finding a basis for a vector space. A basis of a subspace $U = \text{span}(\mathbf{x}_1, \dots, \mathbf{x}_m) \subseteq \mathbb{R}^n$ can be found by executing the following steps:

- ▶ Write the spanning vectors as columns of a matrix \mathbf{A} .
- ▶ Determine the row-echelon form of \mathbf{A} .
- ▶ The spanning vectors associated with the pivot columns are a basis of U .

- ▶ Consider the vector subspace $U \subseteq \mathbb{R}^5$ which is spanned by the following vectors:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix}$$

- ▶ We would like to check if the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ constitute a basis for the subspace U .



Theorem: For a given matrix $\mathbf{A}_{m \times n}$, we have

$$\text{rank}(\mathbf{A}) + \dim(\mathbf{N}(\mathbf{A})) = n$$

. Some consequences of this theorem are

- ▶ For a square matrix \mathbf{A} , if $\dim(\mathbf{N}(\mathbf{A})) = 0$, then \mathbf{A} has *full rank* n .
- ▶ For the general case with n being odd, $\text{rank}(\mathbf{A})$ and $\dim(\mathbf{N}(\mathbf{A}))$ cannot be equal. (Why?)
- ▶ What could be the maximum value of $\dim(\mathbf{N}(\mathbf{A}))$ and can such an \mathbf{A} be characterized?