A Few Notes on Linear Algebra

David Meyer

dmm613@gmail.com

Last update: May 24, 2023

1 Introduction

2 Linear Subspaces

Definition 2.1. Linear Subspace: Let $U \subseteq \mathbb{R}^n$ with $U \neq \emptyset$. Then U is called a *linear subspace* of $\mathbb{R}^n$ if U is closed under linear combination [3]. More specifically, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)}$ be vectors in U and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be scalars in $\mathbb{R}$. Then for U to be a linear subspace we need that

$$\sum_{j=1}^{k} \lambda_j \mathbf{u}^{(j)} \in U$$

That is, all linear combinations of vectors in U and scalars in $\mathbb{R}$ are also in U.

2.1 Characteristics of Linear Subspaces

$U \subseteq \mathbb{R}^n$ is a (linear) subspace iff

(a) $\mathbf{0} \in U$

(b) $\mathbf{u} \in U$ and $\lambda \in \mathbb{R} \Rightarrow \lambda \mathbf{u} \in U$

(c) $\mathbf{u}, \mathbf{v} \in U \Rightarrow \mathbf{u} + \mathbf{v} \in U$

For example, consider a line $U \subseteq \mathbb{R}^2$. Then for all $\mathbf{u}, \mathbf{v} \in U$ and $\lambda \in \mathbb{R}$ we have $\lambda \mathbf{u} \in U$ and $\mathbf{u} + \mathbf{v} \in U$. Thus U is a linear subspace of $\mathbb{R}^2$. This is shown (hopefully) in Figure 1.

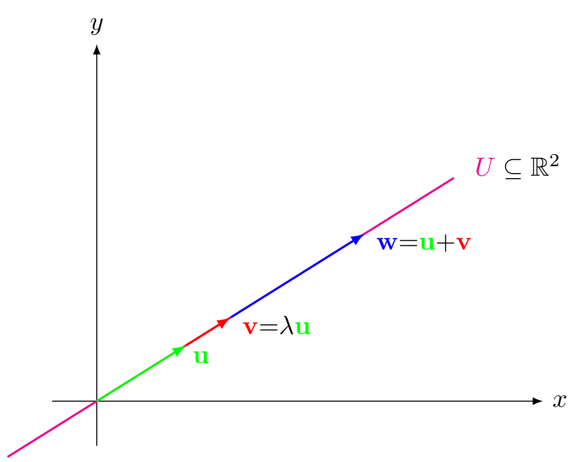

Figure 1: The line U is a linear subspace of $\mathbb{R}^2$

3 Linear Maps

In linear algebra a linear map is a mapping $\mathbf{V} \rightarrow \mathbf{W}$ between two vector spaces which preserves the operations of vector addition and scalar multiplication. Interestingly, the same names and the same definition are also used for the more general case of modules over a ring (need module homomorphism cite).

Definition 3.1. Linear Map: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called *linear* if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$:

(a) $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ # vector addition is *distributive*

(b) $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ # scalar multiplication is *compatible* (aka homogeneity)

Note that in Definition 3.1 (a) the vector addition on the left-hand side is in $\mathbb{R}^n$ while on the right-hand side the vector addition is in $\mathbb{R}^m$. Similarly, in Definition 3.1 (b) the scalar multiplication on the left-hand side is in $\mathbb{R}^n$ while the scalar multiplication on the right-hand side is in $\mathbb{R}^m$.

3.1 Examples

1. $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x$

   This is linear because

   - $f(x + y) = x + y = f(x) + f(y)$
   - $f(\lambda x) = \lambda x = \lambda f(x)$

2. $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^2$

   This is not linear because $f(3 \cdot 1) = f(3) = 9$ and for f a linear map we have $f(3 \cdot 1) = 3 \cdot f(1) = 3 \cdot 1 = 3$. But $9 \neq 3$ so f is not linear.

3. $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x + 1$

   This is not linear because $f(0 \cdot 1) = 1$ but $0 \cdot f(1) = 0$ and $1 \neq 0$.

   Aside: $f(x) = x + 1$ looks linear but under this definition ((a) and (b) above) f is not linear.

3.2 Relationship between linear maps and group homomorphism

As we saw above, in linear algebra, a linear map is a function between two vector spaces that preserves the vector space structure, i.e., it preserves addition and scalar multiplication. A group homomorphism, on the other hand, is a function between two groups that preserves the group structure, i.e., it preserves the group operation.

3.2.1 Group Homomorphism

Let $(G, \diamond)$ and $(H, \circ)$ be groups. Common usage is to use G to refer to $(G, \diamond)$. Similarly, H will refer to $(H, \circ)$. Then a mapping $\phi : G \rightarrow H$ is called a homomorphism iff

$$\phi(x \diamond y) = \phi(x) \circ \phi(y) \ \ \forall x, y \in G$$

Essentially, a homomorphism $\phi : G \rightarrow H$ is a way of exploring the structure of H by varying G using structure preserving transformations. Specifically, ϕ preserves the group structure.

Example 3.1. Define a map

$$\phi : G \to H$$

where $G = \mathbb{Z}$ and $H = \mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$ is the standard group of order two. Then define $\phi : \mathbb{Z} \to \mathbb{Z}_2$ by the rule

$$\phi(x) = \begin{cases} 0 & \text{if } x \text{ is even} \\ 1 & \text{if } x \text{ is odd} \end{cases}$$

It is not too hard to check that ϕ is a homomorphism. Suppose that x and y are two integers. Then there are four cases:

- x and y are both even

  In this case $\phi(x+y) = 0$ (even + even = even). Here $\phi(x)+\phi(y) = 0+0 = 0$ so $\phi(x+y) = \phi(x)+\phi(y) = 0+0 = 0$.

- x and y are both odd

  In this case $\phi(x + y) = 0$ (odd + odd = even). Here $\phi(x) + \phi(y) = 1 + 1 = 2 \mod 2 = 0$, so $\phi(x + y) = \phi(x) + \phi(y) = 1 + 1 = 2 \mod 2 = 0$.

- x is even and y is odd or x is odd and y is even

  In this case one is even and the other is odd and $x + y$ is odd. Here $\phi(x + y) = 1$ and $\phi(x) + \phi(y) = 1 + 0 = 1$ so $\phi(x + y) = \phi(x) + \phi(y)$.

Thus ϕ is a homomorphism. Note that in this example $\diamond = +$ (normal addition in $\mathbb{Z}$) and $\circ = +$ (addition mod 2 in $\mathbb{Z}_2$).

Not surprisingly, there is a close relationship between linear maps and group homomorphisms because any linear map between two vector spaces can be seen as a group homomorphism between the additive groups of those vector spaces. Specifically, if we have a linear map $f : \mathbf{V} \to \mathbf{W}$ between two vector spaces $\mathbf{V}$ and $\mathbf{W}$ over the same field $\mathbf{F}$, then we can define a group homomorphism from $\phi : (\mathbf{V}, +) \to (\mathbf{W}, +)$ by sending each vector $\mathbf{v} \in \mathbf{V}$ to $f(\mathbf{v}) \in \mathbf{W}$. This is a group homomorphism because f preserves addition (its a linear map, Definition 3.1 (a) above) and hence it preserves the group structure.

On the other hand, any group homomorphism between two additive groups of vector spaces can be seen as a linear map between those vector spaces. Specifically, if we have a group homomorphism $\phi : (\mathbf{V}, +) \to (\mathbf{W}, +)$ between two additive groups of vector spaces, then we can define a linear map from $f : \mathbf{V}$ to $\mathbf{W}$ by sending each vector $\mathbf{v} \in \mathbf{V}$ to $\phi(\mathbf{v}) \in \mathbf{W}$. This is a linear map because ϕ preserves vector addition and scalar multiplication and therefore it preserves the structure of the underlying vector spaces.

So basically, every linear map f between two vector spaces over the same field can be thought of as a group homomorphism ϕ between the additive groups of those vector spaces, and every group homomorphism between two groups can be thought of as a linear map between the vector spaces over the same field that are associated with those groups. So a linear map is a homomorphism of vector spaces. However, if the vector spaces have additional structure, for example a ring, an algebra, or a lie algebra, then a linear map is not always a homomorphism of those additional structures.

4 Matrices Induce Linear Maps

Consider a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{x} \in \mathbb{R}^n$, and define $f_{\boldsymbol{A}} : \mathbb{R}^n \to \mathbb{R}^m$ by $\mathbf{x} \mapsto \boldsymbol{A}\mathbf{x}$.

Proposition 4.1. $f_{\boldsymbol{A}}$ is a linear map

Ok, but why? Here is one way to look at it: $f_{\boldsymbol{A}}$ a linear map means

- $f_{\boldsymbol{A}}(\mathbf{x}+\mathbf{y}) = f_{\boldsymbol{A}}(\mathbf{x}) + f_{\boldsymbol{A}}(\mathbf{y}) \Rightarrow \boldsymbol{A}(\mathbf{x}+\mathbf{y}) = \boldsymbol{A}\mathbf{x} + \boldsymbol{A}\mathbf{y}$ # $f_{\boldsymbol{A}}$ is distributive (Definition 3.1 (a))
- $f_{\boldsymbol{A}}(\lambda\mathbf{x}) = \lambda f_{\boldsymbol{A}}(\mathbf{x}) \Rightarrow \boldsymbol{A}(\lambda\mathbf{x}) = \lambda\boldsymbol{A}\mathbf{x}$ # $f_{\boldsymbol{A}}$ is compatible (Definition 3.1 (b))

Here's an example: Let $\boldsymbol{A} \in \mathbb{R}^{m\times 2}$ with $\boldsymbol{A} = \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}$ and let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ be vectors in $\mathbb{R}^2$.

Then

$$
\begin{aligned}
\boldsymbol{A}(\mathbf{x}+\mathbf{y}) &= \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) && \text{\# definitions of } \boldsymbol{A}, \mathbf{x}, \text{ and } \mathbf{y} \\
&= \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} && \text{\# matrix (vector) addition in } \mathbb{R}^2 \\
&= \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix}\begin{pmatrix} x_1 + y_1 \end{pmatrix} + \begin{pmatrix} | \\ a_2 \\ | \end{pmatrix}\begin{pmatrix} x_2 + y_2 \end{pmatrix} && \text{\# definition of the matrix product} \\
&= \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} x_1 + \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} y_1 + \begin{pmatrix} | \\ a_2 \\ | \end{pmatrix} x_2 + \begin{pmatrix} | \\ a_2 \\ | \end{pmatrix} y_2 && \text{\# matrix product distributes over addition} \\
&= \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} x_1 + \begin{pmatrix} | \\ a_2 \\ | \end{pmatrix} x_2 + \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} y_1 + \begin{pmatrix} | \\ a_2 \\ | \end{pmatrix} y_2 && \text{\# vector addition is commutative} \\
&= \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} && \text{\# definition of the matrix product} \\
&= \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\boldsymbol{x} + \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}\boldsymbol{y} && \text{\# } \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and } \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \boldsymbol{A}\mathbf{x} + \boldsymbol{A}\mathbf{y} && \text{\# } \boldsymbol{A} = \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}
\end{aligned}
$$

So $\boldsymbol{A}(\mathbf{x}+\mathbf{y}) = \boldsymbol{A}\mathbf{x} + \boldsymbol{A}\mathbf{y}$. Here we see that the matrix $\boldsymbol{A}$, which is just a table of real (or complex) numbers, induces the abstract linear map $f_{\boldsymbol{A}}$ (and vice versa) [1].

Suppose f is a linear map and let $\boldsymbol{x} = \begin{pmatrix} x_1, x_2, \dots, x_n \end{pmatrix}$ be a vector in $\mathbb{R}^n$. Then define the canonical unit vectors in $\mathbb{R}^n$ as follows:

Definition 4.1. Canonical Unit Vectors: The canonical unit vectors in $\mathbb{R}^n$ are the vectors $\mathbf{e}_i$

$$B = \{\mathbf{e}_i : 1 \leq i \leq n\}$$

where $\mathbf{e}_i$ is the i^{th} unit vector, that is, it has a one in the i^{th} coordinate (position) and zeros everywhere else[1]. The set B forms a basis, sometimes called the canonical basis, for the vector space $\mathbb{R}^n$.

The basis vectors $\mathbf{e}_i$ have column vector format

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \; \mathbf{e}_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \; \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Here is an interesting observation about linear maps and basis vectors:

$$\begin{aligned} f(\boldsymbol{x}) &= f(x_1\boldsymbol{e_1} + x_2\boldsymbol{e_2} + \ldots + x_n\boldsymbol{e_n}) && \#\ \boldsymbol{x} \text{ is a linear combination of basis vectors} \\ &= f(x_1\boldsymbol{e_1}) + f(x_2\boldsymbol{e_2}) + \ldots + f(x_n\boldsymbol{e_n}) && \#\ f \text{ a linear map} \Rightarrow f \text{ is distributive (Definition 3.1 (a))} \\ &= x_1 f(\boldsymbol{e_1}) + x_2 f(\boldsymbol{e_2}) + \ldots + x_n f(\boldsymbol{e_n}) && \#\ f \text{ a linear map} \Rightarrow f \text{ is compatible (Definition 3.1 (b))} \end{aligned}$$

Since $f(\boldsymbol{x}) = x_1 f(\boldsymbol{e_1}) + x_2 f(\boldsymbol{e_2}) + \ldots + x_n f(\boldsymbol{e_n})$ apparently in order to understand f you only have to know how f maps the basis vectors.

5 Conclusions

6 Acknowledgements

LaTeX Source

`https://www.overleaf.com/read/pmsddwthrchw`

References

[1] Wikipedia Contributors. Linear Map — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Linear_map&oldid=1024503856`, 2021. [Online; accessed 6-June-2021].

[2] Wikipedia Contributors. One-hot — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=One-hot&oldid=1127084325`, 2022. [Online; accessed 17-April-2023].

[3] Wikipedia Contributors. Linear Subspace — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Linear_subspace&oldid=1147122525`, 2023. [Online; accessed 21-April-2023].

[1] The canonical unit vectors look surprisingly like the one-hot encoding used in machine learning [2].

Appendix A: Draw Some Matrices

$$A^{n\times m} = \begin{pmatrix} — & a_1 & — \\ — & a_2 & — \\ & \vdots & \\ — & a_n & — \end{pmatrix}$$

$$A^T = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \dots & a_n \\ | & | & & | \end{pmatrix}$$

$$B^{n\times k} = \begin{pmatrix} — & b_1 & — \\ — & b_2 & — \\ & \vdots & \\ — & b_n & — \end{pmatrix}$$

$$\begin{aligned} AB := A^T B &= \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \dots & a_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} — & b_1 & — \\ — & b_2 & — \\ & \vdots & \\ — & b_n & — \end{pmatrix} \\ &= \begin{pmatrix} a_1^T b_1 & a_1^T b_2 & \dots & a_1^T b_k \\ a_2^T b_1 & a_2^T b_2 & \dots & a_2^T b_k \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \dots & a_m^T b_k \end{pmatrix} \\ &= \begin{pmatrix} \langle a_1, b_1\rangle & \langle a_1, b_2\rangle & \dots & \langle a_1, b_k\rangle \\ \langle a_2, b_1\rangle & \langle a_2, b_2\rangle & \dots & \langle a_2, b_k\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_m, b_1\rangle & \langle a_m, b_2\rangle & \dots & \langle a_m, b_k\rangle \end{pmatrix} \end{aligned}$$

$$A^T \quad = \quad \begin{pmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_n^T & — \end{pmatrix}$$