Coursera – Coding the Matrix notes

Alejandro Anzola Ávila

February 15, 2020

Chapter 0

Week 0: The function and the field

0.1 The function and other preliminaries

0.1.1 Set terminology and notation

Definition 0.1.1 (Set)**.** *A set is an unordered collection of objects.*

- $\in$: indicates that an object belongs to a set. (e.g. $a \in \{a, b, \ldots\}$)
- $\mathbb{A} \subseteq \mathbb{B}$: "$\mathbb{A}$ is a **subset** of $\mathbb{B}$". Every element of $\mathbb{A}$ is also an element of $\mathbb{B}$
- $\mathbb{A} = \mathbb{B}$: two sets are equal if they contain exactly the same elements.

Set expressions

$\{x \in \mathbb{R} : x \geq 0\}$ is the set of nonnegative numbers. First part specifies where the elements of the set comes from and introduces variables. The second part gives a rule that restricts which elements specified in the first part actually get to make it into the set.

Definition 0.1.2 (Cardinality)**.** *If a set $\mathbb{S}$ is not infinite, we use $|\mathbb{S}|$ to denote the number of elements or* cardinality *of the set.*

Definition 0.1.3. $\mathbb{A} \times \mathbb{B}$ *is the set of all pairs (a, b) where $a \in \mathbb{A}$ and $b \in \mathbb{B}$*

0.1.2 The function

Informally, for each input element in a set $\mathbb{A}$, a function assigns a single output element from another set $\mathbb{B}$

- $\mathbb{A}$ is called the **domain** of the function
- $\mathbb{B}$ is called the **co-domain**

Definition 0.1.4 (Function)**.** *A function is a set of pairs (a, b) no two of which have the same first element.*

Definition 0.1.5 (Image)**.** *The output of a given input is called the* image *of that input. The image of q under a function f is denoted $f(q)$*

If $f(q) = r$, we say q maps to r under f. In Mathese, we write this as $q \mapsto r$.

The set from which all the outputs are chosen is called the co-domain. We write:

$$f : \mathbb{D} \rightarrow \mathbb{F}$$

when we want to say that f is a function with domain $\mathbb{D}$ and co-domain $\mathbb{F}$.

Definition 0.1.6 (Image of a function)**.** *The image of a function is the set of all images of inputs. Mathese:* $\operatorname{Im} f$

Example. $\cos : \mathbb{R} \rightarrow \mathbb{R}$, *which means the domain is $\mathbb{R}$, and the co-domain is $\mathbb{R}$. The image of* $\cos(x)$, $\operatorname{Im}\cos$ *is* $\{x \in \mathbb{R} : -1 \leq x \leq 1\}$.

Definition 0.1.7. *For sets $\mathbb{F}$ and $\mathbb{D}$, $\mathbb{F}^{\mathbb{D}}$ denotes all functions from $\mathbb{D}$ to $\mathbb{F}$.*

Proposition 0.1.7.1. *For finite sets, $|\mathbb{F}^{\mathbb{D}}| = |\mathbb{F}|^{|\mathbb{D}|}$.*

Definition 0.1.8 (Identity function)**.** *For any domain $\mathbb{D}$. $\mathrm{id}_{\mathbb{D}} : \mathbb{D} \rightarrow \mathbb{D}$ maps each domain element d to itself.*

Definition 0.1.9 (Functional composition)**.** *For functions $f : \mathbb{A} \rightarrow \mathbb{B}$ and $g : \mathbb{B} \rightarrow \mathbb{C}$, the functional composition of f and g is the function $(g \circ f) : \mathbb{A} \rightarrow \mathbb{C}$ defined by $(g \circ f)(x) = g(f(x))$.*

Proposition 0.1.9.1. $h \circ (g \circ f) = (h \circ g) \circ f$

Definition 0.1.10 (Functional inverses)**.** *Functions f and g are functional inverses if $f \circ g$ and $g \circ f$ are defined and are identity functions. A function that has an inverse is invertible.*

Definition 0.1.11. *$f : \mathbb{D} \to \mathbb{F}$ is* ***one-to-one*** *if $f(x) = f(y)$ implies $x = y$.*

Definition 0.1.12. *$f : \mathbb{D} \to \mathbb{F}$ is* ***ontox*** *if for every $z \in \mathbb{F}$ there exists an a such that $f(a) = z$.*

Proposition 0.1.12.1. *Invertible functions are one-to-one.*

Theorem 0.1.1 (Function Invertibility Theorem)**.** *A function f is invertible if and only if it is one-to-one and onto.*

0.2 The Field: Introduction to complex numbers

$i = \sqrt{-1}$ is an imaginary number, this is a solution to an equation such as $x^2 = -1$. For $(x-1)^2 = 9$, the solution is $x = 1 + 3i$.

A **complex number** has a real part and an imaginary part.

0.2.1 Field notation

When we want to refer to a field without specifying which field we will use the notation $\mathcal{F}$.

We study three fields;

- The field $\mathbb{R}$ of real numbers.
- The field $\mathbb{C}$ of complex numbers.
- The finite field $GF(2)$, which consists of 0 and 1 under $\mod 2$ arithmetic.

0.3 The Field of playing with $\mathbb{C}$

We can interpret real and imaginary parts of a complex number as x and y coordinates. Assume that $z \in \mathbb{C}$.

- **Translation** $f(z) = z + z_0, z_0 \in \mathbb{C}$. A translation can "move" a picture anywhere in the complex plane.
- **Scaling** $f(z) = mz, m \in \mathbb{R}$.
- **Invert** $f(z) = (-1)z$.
- **Rotate counterclockwise by 90 degreesx** $f(z) = iz$.
- **Rotating by an angle** $f(z) = z \cdot e^{\tau i}$, does rotation by angle τ.

0.4 The Field of playing with $GF(2)$

$GF(2)$ = Galois Field 2, has just two elements: 0 and 1.

- Addition is like exclusive-or. (e.g. $\mathrm{XOR}(a, b) = a \not\equiv b; a, b \in GF(2)$)
- Multiplication is just like normal multiplication.

Chapter 1

Week 1: The vector

1.1 What is a vector?

A vector is an array of d numbers, and also can be thought as functions that maps from $\{0, 1, \ldots, d-1\}$ to $\mathcal{F}$ with $\mathcal{F}^d$ as the notation.

A vector most of whose values are zero is called a *sparse* vector. If no more than k of the entries are nonzero, we say the vector is k-sparse. A k-sparce vector can be represented using space proportional to k.

1.2 Vector addition and scalar-vector multiplication

Definition 1.2.1 (Vector addition)**.**

$$[u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n] = [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n]$$

Definition 1.2.2 (Zero vector)**.** *The D-vector whose entries are all zero is the zero vector, written* $\mathbf{0}_D$ *or just* $\mathbf{0}$.

$$\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$$

Definition 1.2.3 (Associativity)**.**

$$(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$$

Definition 1.2.4 (Commutativity)**.**

$$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$$

For vectors, we refer to field elements as scalars, we use them to scale vectors: $\alpha\boldsymbol{v}$. Greek letters (e.g. α, β, γ) denote scalars.

Definition 1.2.5. *Multiplying a vector $\boldsymbol{v}$ by a scalar α is defined as multiplying each entry of $\boldsymbol{v}$ by α.*

$$\alpha[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] = [\alpha\boldsymbol{v}_1, \ldots, \alpha\boldsymbol{v}_n]$$

The set of points $\{\alpha\boldsymbol{v} : \alpha \in \mathbb{R}\}$ forms the line through the origin and $\boldsymbol{v}$.

An expression of the form $\alpha\boldsymbol{u} + \beta\boldsymbol{v}$ where $0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$, and $\alpha + \beta = 1$ is called a *convex combination* of $\boldsymbol{u}$ and $\boldsymbol{v}$. An expression of the form $\alpha\boldsymbol{u} + \beta\boldsymbol{v}$ where $\alpha + \beta = 1$ is called and *affine combination* of $\boldsymbol{u}$ and $\boldsymbol{v}$.

1.3 Dot-product

Definition 1.3.1 (Dot-product of two D-vectors)**.** *Dot-product of two D-vectors is the sum of product of corresponding entries:*

$$\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{k \in \{1,\ldots,D\}} \boldsymbol{u}_k \boldsymbol{v}_k$$

1.3.1 Linear equations

Definition 1.3.2. *A linear equation is an equation of the form*

$$\boldsymbol{a} \cdot \boldsymbol{x} = \beta$$

where $\boldsymbol{a}$ is a vector, β is a scalar, and $\boldsymbol{x}$ is a vector of variables.

Algebraic properties of dot-product:

- Commutativity $\boldsymbol{v} \cdot \boldsymbol{x} = \boldsymbol{x} \cdot \boldsymbol{v}$
- Homogeneity $(\alpha\boldsymbol{u}) \cdot \boldsymbol{v} = \alpha(\boldsymbol{u} \cdot \boldsymbol{v})$
- Distributive law $(\boldsymbol{v}_1 + \boldsymbol{v}_2) \cdot \boldsymbol{x} = \boldsymbol{v}_1 \cdot \boldsymbol{x} + \boldsymbol{v}_2 \cdot \boldsymbol{x}$

Chapter 2

Week 2: The vector space

2.1 Linear combinations

Definition 2.1.1. *An expression*

$$\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$$

is a linear combination of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$. The scalars $\alpha_1, \ldots, \alpha_n$ are the coefficients of the linear combination.

2.2 Span

Definition 2.2.1 (Span)**.** *The set of all linear combinations of some vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is called the span of these vectors. Written $\mathrm{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$.*

Definition 2.2.2. *Let $\mathbb{V}$ be a set of vectors. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are vectors such that $\mathbb{V} = \mathrm{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, then*

- *We say $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a generating set for $\mathbb{V}$.*
- *We refer to the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ as generators for $\mathbb{V}$.*

If we use the vectors $[1, 0, 0]$, $[0, 1, 0]$, and $[0, 0, 1]$:

$$[x, y, z] = x[1, 0, 0] + y[0, 1, 0] + z[0, 0, 1]$$

These are called *standard generators* for $\mathbb{R}^3$.

2.3 Geometry of sets of vectors

Span of a single nonzero vector $\boldsymbol{v}$:

$$\mathrm{Span}\{\boldsymbol{v}\} = \{\alpha\boldsymbol{v} : \alpha \in \mathbb{R}\}$$

This is the line through the origin and $\boldsymbol{v}$, One-dimensional. The span of the empty set: just the origin, Zero-dimensional.

One way to specify a plane can be with:

$$\{(x, y, z) : ax + by + cz = 0\}$$

Using dot-product, we could rewrite as

$$\{[x, y, z] : [a, b, c] \cdot [x, y, z] = 0\}$$

We can specify a line in three dimensions:

$$\{[x, y, z] : \boldsymbol{a}_1 \cdot [x, y, z] = 0, \boldsymbol{a}_2 \cdot [x, y, z] = 0\}$$

Two ways to represent a geometric object (line, plane, etc.) containing the origin:

- Span of some vectors.
- Solution set of some system of linear equations with zero right hand sides.

What is common with these two representations? R: Subset $\mathcal{F}^D$ satisfies three properties:

1. Subset contains the zero vector $\mathbf{0}$.
2. If subset contains $\boldsymbol{v}$ then it contains $\alpha\boldsymbol{v}$ for every scalar α.
3. If subset contains $\boldsymbol{u}$ and $\boldsymbol{v}$ then it contains $\boldsymbol{u} + \boldsymbol{v}$.

Definition 2.3.1 (Vector space)**.** *Any subset $\mathbb{V}$ of $\mathcal{F}^D$ satisfying the three properties is called a vector space.*

Definition 2.3.2. *If $\mathbb{U}$ is also a vector space and $\mathbb{U}$ is a subset of $\mathbb{V}$ then $\mathbb{U}$ is called a subspace of $\mathbb{V}$.*

2.3.1 Convex hull

Definition 2.3.3. *For vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ over $\mathbb{R}$, a linear combination*

$$\alpha_1 \boldsymbol{v}_1, \ldots, \alpha_n \boldsymbol{v}_n$$

is a convex combination if the coefficients are all nonnegative and they sum to 1*.*

2.4 Vector spaces

To represent an object that doesn't contain the origin we sum a vector $\boldsymbol{c}$ and redefine the definition as:

$$\{\boldsymbol{c} + \boldsymbol{v} : \boldsymbol{v} \in \mathbb{V}\}$$

It can also be abbreviated as $\{\boldsymbol{c} + \mathbb{V}\}$.

Definition 2.4.1. *A linear combination*

$$\alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 + \cdots + \alpha_n \boldsymbol{u}_n$$

where

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$$

is an affine combination.

Definition 2.4.2. *The set of all affine combinations of vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ is called the* affine hull *of those vectors.*

Affine hull of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n = \{\boldsymbol{u}_1 + \mathrm{Span}\{\boldsymbol{u}_2 - \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n - \boldsymbol{u}_1\}\}$. This shows that the affine hull is an affine space.

In general, a geometric object can be expressed as the solution set of a system of linear equations.

$$\{\boldsymbol{x} : \boldsymbol{a}_1 \cdot \boldsymbol{x} = \beta_1, \cdots, \boldsymbol{a}_m \cdot \boldsymbol{x} = \beta_m\}$$

Conversely, is the solution set an affine space? Consider solution set of a contradictory system of equations, e.g. $1x = 1, 2x = 1$:

- Solution set is empty
- but a vector space $\mathbb{V}$ always contains the zero vector,
- so an affine space $\boldsymbol{u}_1 + \mathbb{V}$ always contains at least one vector.

Turns out this is the only exception:

Theorem 2.4.1. *The solution set of a linear system is either empty or an affine space.*

Definition 2.4.3. *A linear equation* $\boldsymbol{a} \cdot \boldsymbol{x} = 0$ *with zero right-hand side is a homogeneous linear equation. A system of homogeneous linear equations is called a homogeneous linear system.*

Lemma 2.4.2. *Let* $\boldsymbol{u}_1$ *be a solution to a linear system. Then, for any other vector* $\boldsymbol{u}_2$*,* $\boldsymbol{u}_2$ *is also a solution if and only if* $\boldsymbol{u}_2 - \boldsymbol{u}_1$ *is a solution to the corresponding homogeneous linear system.*

Chapter 3

Week 3: The matrix

3.1 What is a matrix

The traditional notion of a matrix: a two dimensional array

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

where $\boldsymbol{A}$ is a 2×3 matrix.

For a matrix $\boldsymbol{A}$, the i, j element of $\boldsymbol{A}$ is the element in row i, column j. It is traditionally written as $\boldsymbol{A}_{i,j}$

Definition 3.1.1. *For finite sets $\mathbb{R}$ and $\mathbb{C}$, and $\mathbb{R} \times \mathbb{C}$ matrix over field $\mathcal{F}$ is a function from $\mathbb{R} \times \mathbb{C}$ to $\mathcal{F}$.*

Definition 3.1.2. *$D \times D$ identity matrix is the matrix $\mathbb{1}_D$ such that $\mathbb{1}_D[k, k] = 1$ for all $k \in \{1, , \ldots, D\}$ and zero elsewhere.*

Usually we omit the subscript when D is clear from the context. Often letter $\boldsymbol{I}$ is used instead of $\mathbb{1}$

3.1.1 Column space and row space

One simple role for a matrix: packing together a bunch of columns or rows.

Definition 3.1.3. *Two vectors associated with a matrix $\boldsymbol{M}$:*

- *Column space of $\boldsymbol{M} = \mathrm{Span}\{$columns of $\boldsymbol{M}\}$. Written $\mathrm{Col}\boldsymbol{M}$.*
- *Row space of $\boldsymbol{M} = \mathrm{Span}\{$rows of $\boldsymbol{M}\}$. Written $\mathrm{Row}\boldsymbol{M}$.*

3.1.2 Transpose

Transpose swaps rows and columns. It is written as $\boldsymbol{A}^\top$.

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

$$\boldsymbol{A}^\top = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$$

3.1.3 Matrices as vectors

A matrix can be interpreted as a vector:

- an $\mathbb{R} \times \mathbb{S}$ matrix is a function from $\mathbb{R} \times \mathbb{S}$ to $\mathcal{F}$,
- it can be interpreted as and $\mathbb{R} \times \mathbb{S}$–vector:
  - scalar-vector multiplication
  - vector addition