Mathematical Toolkit **Autumn 2015**

Lecture 2: September 30, 2015

Lecturer: Madhur Tulsiani

1 Linear Independence and Bases

We recall the definition of a basis and the Steinitz exchange princliple from the previous lecture.

Definition 1.1 *A set B is said to be a basis for the vector space V if B is linearly independent and $Span(B) = V$.*

We emphasize again that the definition of the span only involves linear combinations of finitely many elements. A basis such as above is known as a Hamel basis.

Proposition 1.2 (Steinitz exchange principle) *Let $\{v_1, \ldots, v_k\}$ be linearly independent and $\{v_1, \ldots, v_k\} \subseteq Span(\{w_1, \ldots, w_n\})$. Then $\forall i \in [k]$ $\exists j \in [n]$ such that $\{v_1, \ldots, v_k\} \setminus \{v_i\} \cup \{w_j\}$ is linearly independent.*

A vector space V is said to be finitely generated if there exists a finite set S such that Span $(S) = V$. It is easy to see that a finitely generated vector space has a basis (which is a subset of the generating set S). Also, the following is an easy corollary of the Steinitz exchange principle.

Corollary 1.3 *All bases of a finitely generated vector space have equal size.*

To prove the existence of a basis for every vector space, we will need Zorn's lemma (which is equivalent to the axiom of choice). We first define the concepts needed to state and apply the lemma.

Definition 1.4 *Let X be a non-empty set. A relation $\preceq$ between elements of X is called a* partial order

- $x \preceq x$ *for all* $x \in X$.
- $x \preceq y, y \preceq x \Rightarrow x = y$.
- $x \preceq y, y \preceq z \Rightarrow x \preceq z$.

The relation is called a partial order since not all the elements of X may be realted. A subset $S \subseteq X$ is called totally ordered *if for every $x, y \in S$ we have $x \preceq y$ or $y \preceq x$. A set $S \subseteq X$ is called* bounded *if there exists $x_0 \in X$ such that $x \preceq x_0$ for all $x \in S$. An element $x_0 \in X$ is maximal if there does not exist any other $x \in X$ such that $x_0 \preceq x$.*

Proposition 1.5 (Zorn's lemma) *Let X be a partially ordered set such that every totally ordered subset of X is bounded. Then X contains a maximal element.*

We can use Zorn's lemma to in fact prove a stronger statement than the existence of a basis.

Proposition 1.6 *Let V be a vector space over a field $\mathbb{F}$ and let S be a linearly independent subset. Then there exists a basis B of V containing the set S.*

Proof: Let X be the set of all linearly independent subsets of V that contain S. For $S_1, S_2 \in X$, we say that $S_1 \preceq S_2$ if $S_1 \subseteq S_2$. Let Y be a totally ordered subset of X. Define S_0 as

$$S_0 \ := \ \cup_{T \in Y} T \ = \ \{v \in V \mid \exists T \in Y \text{ such that } v \in T\} \ .$$

Then we claim that S_0 is linearly independent and is hence in X. Is is clear that $T \preceq S_0$ for all $T \in Y$ and this will prove that Y is bounded by T. By Zorn's lemma this shows that X contains a maximal element (say) B, which must be a basis containing S.

To show that S_0 is linearlly independent, note that we only need to show that no *finite* subset of S_0 is linearly dependent. Indeed, let $\{v_1, \ldots, v_k\}$ be a finite linearly subset of S_0. By the definition of S_0, there exists a $T \in X$ such that $\{v_1, \ldots, v_k\} \subseteq T$. Thus, $\{v_1, \ldots, v_k\}$ must be linearly indepdent. This proves the claim. ■

Lagrange Interpolation

Lagrange interpolation is used to find the unique polynomial of degree at most $n-1$, taking given values at n distinct points. We can derive the formula for such a polynomial using basic linear algebra.

Let $a_1, \ldots, a_n \in \mathbb{R}$ be distinct. Say we want to find the unique (why?) polynomial p of degree at most $n-1$ satisfying $p(a_i) = b_i \ \forall i \in [n]$. Recall that the space of polynomials of degree at most $n-1$ with real coefficients, denoted by $\mathbb{R}^{\leq n-1}[x]$, is a vector space. Also, recall from the last lecture that if we define $g(x)$ as $\prod_{i=1}^{n}(x - a_i)$, the degree $n-1$ polynomials defined as

$$f_i(x) \ = \ \frac{g(x)}{x - a_i} \ = \ \prod_{j \neq i}^{n} (x - a_j) \, ,$$

are n linearly independent polynomials in $\mathbb{R}^{\leq n-1}[x]$. Thus, they must form a basis for $\mathbb{R}^{\leq n-1}[x]$ and we can write the prequired polynomial, say p as

$$p \ = \ \sum_{i=1}^{n} c_i \cdot f_i \, ,$$

for some $c_1, \ldots, c_n \in \mathbb{R}$. Evaluating both sides at a_i gives $p(a_i) = b_i = c_i \cdot f_i(a_i)$. Thus, we get

$$p(x) \ = \ \sum_{i=1}^{n} \frac{b_i}{f_i(a_i)} \cdot f_i(x) \, .$$

2 Linear Transformations

Definition 2.1 *Let V and W be vector spaces over the same field $\mathbb{F}$. A map $\varphi : V \to W$ is called a* linear transformation *if*

- $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \quad \forall v_1, v_2 \in V$.
- $\varphi(c \cdot v) = c \cdot \varphi(v) \quad \forall v \in V$.

Example 2.2 *The following are all linear transformations:*

- *A matrix $A \in \mathbb{R}^{m \times n}$ defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.*
- *$\varphi : C([0,1], \mathbb{R}) \to C([0,1], \mathbb{R})$ defined by $\varphi(f)(x) = f(1-x)$.*
- *$\varphi_{left} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ defined by $\varphi_{left}(f)(n) = f(n+1)$.*
- *The derivative operator acting on $\mathbb{R}[x]$.*

Proposition 2.3 *Let V, W be vector spaces over $\mathbb{F}$ and let B be a basis for V. Let $\alpha : B \to W$ be an arbitrary map. Then there exists a unique linear transformation $\varphi : V \to W$ satisfying $\varphi(v) = \alpha(v)$ $\forall v \in B$.*

Definition 2.4 *Let $\varphi : V \to W$ be a linear transformation. We define its* kernel *and* image *as:*

- $ker(\varphi) := \{v \in V \mid \varphi(v) = 0_W\}$.
- $im(\varphi) = \{\varphi(v) \mid v \in V\}$.

Proposition 2.5 *$ker(\varphi)$ is a subspace of V and $im(\varphi)$ is a subspace of W.*

Proposition 2.6 (rank-nullity theorem) *If V is a finite dimesional vector space and $\varphi : V \to W$ is a linear transformation, then*

$$dim(ker(\varphi)) + dim(im(\varphi)) = dim(V)\,.$$

$\dim(\mathrm{im}(\varphi))$ is called the rank and $\dim(\ker(\varphi))$ is called the nullity of φ.

3 Answer to the puzzle problem

In the previous lecture, we asked the following question:

Problem 3.1 ([Mat10]) *Let x be an irrational number. Use linear algebra to show that a rectangle with sides 1 and x cannot be tiled with a finite number of non-overlapping squares.*

We can now solve it given our current knowedge of linear algebra. Recall that $\mathbb{R}$ is a vector space over $\mathbb{Q}$ and 1 and x are linearly independent elements of this vector space. Let us assume that $S_1, \ldots, S_n$ are squares with side lengths $l_1, \ldots, l_n$, which tile the rectangle R. Let $S = \text{Span}\left(\{1, x, l_1, \ldots, l_n\}\right)$.

Since there exists a basis for S containing 1 and x, and since any map from this basis to $\mathbb{R}$ defines a unique linear transformation, there exists a linear transformation $\varphi : S \to \mathbb{R}$ satisfying $\varphi(1) = 1$ and $\varphi(x) = -1$. Define the (area like) function $\mu : S \times S \to \mathbb{R}$ as $\mu(a, b) = \varphi(a) \cdot \varphi(b)$. For a rectangle R_0 with sides $a, b \in S$, we use $\mu(R_0)$ to denote $\mu(a, b)$.

One can show that if we extend all line segments bounding the squares to the sides of R then the sides of all new rectangles generated this way, lie in S and hence μ is defined for all these rectangles. Also, it is easy to check that μ adds like area i.e., if a rectangle R_3 is split in to R_1 and R_2, then $\mu(R_3) = \mu(R_1) + \mu(R_2)$. This gives

$$\varphi(1) \cdot \varphi(x) \;=\; \mu(R) \;=\; \sum_{i=1}^{n} \mu(S_i) \;=\; \sum_{i=1}^{n} \left(\varphi(l_i)\right)^2 ,$$

which is a contradiction since the LHS is -1 while the RHS is non-negative.

References

[Mat10] Jiří Matoušek, *Thirty-three miniatures: Mathematical and algorithmic applications of linear algebra*, vol. 53, American Mathematical Soc., 2010.