II.B. Basis and dimension

How would you explain that a plane has two dimensions? Well, you can go in two independent directions, and no more. To make this idea precise, we formulate the

II.B.1. Definition. A (finite) collection \{\vec{v}_1, \ldots, \vec{v}_m\} in a vector space \(V\) (over a field \(F\)) is a basis of \(V\) if (a) they are linearly independent (over \(F\)), and (b) they span \(V\) (over \(F\)).

The “over \(F\)” means that the linear combinations are understood to be of the form \(\sum \alpha_i \vec{v}_i\), with \(\alpha_i \in F\). As usual, our “default” will be \(F = \mathbb{R}\).

II.B.2. Example. A basis of \(\mathbb{R}^m\) is given by the standard basis vectors \(\{\hat{e}_i\}_{i=1}^m\).

Of course, one could say the same for any field: the standard basis vectors are also a basis for \(\mathbb{C}^m\), viewed as a vector space over \(\mathbb{C}\). But what if we view \(\mathbb{C}^m\) as a vector space over \(\mathbb{R}\)? Then you need the \(2m\) vectors \(\hat{e}_1, \ldots, \hat{e}_m\) and \(i\hat{e}_1, \ldots, i\hat{e}_m\)!

II.B.3. Example. Let \(P_2(x, y)\) denote the vector space of quadratic polynomials in two variables \(x\) and \(y\), with real coefficients. This has the “monomial basis” \(\{\rho_i^2(x, y)\}_{i=1}^6 = \{1, x, y, xy, x^2, y^2\}\). Similarly, \(\{\rho_i^3(x, y)\}_{i=1}^{10} = \{1, x, y, xy, x^2, y^2, x^2y, y^2x, x^3, y^3\}\) gives a basis for the space \(P_3(x, y)\) of cubic polynomials.

We’ll give a geometric application of this example below.

II.B.4. Example. Let \(A\) be an \(m \times n\) matrix, and \(V\) be the vector subspace of \(\mathbb{R}^n\) consisting of solutions to \(Ax = 0\). How do you compute a basis of \(V\)?

\(^3\)This is not to say that you can’t have infinite bases; indeed, non-finitely-generated vector spaces (like \(P_\infty(x) := \) all polynomials) may have such a basis. We’re just not defining or discussing them at this stage.

\(^4\)By “quadratic” resp. “cubic” polynomials, we mean polynomials of degree \(\leq 2\) resp. \(\leq 3\). (We include lower-degree polynomials to have closure under addition.)
By making use of the row-reduction algorithm, of course! For instance, we might have

$$A = \begin{pmatrix}
0 & 0 & 1 & -1 & -1 \\
2 & 4 & 2 & 4 & 2 \\
2 & 4 & 3 & 3 & 3 \\
3 & 6 & 6 & 3 & 6 \\
\end{pmatrix} \rightarrow \text{rref}(A) = \begin{pmatrix}
1 & 2 & 0 & 3 & 0 \\
0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}.$$  

The free variables are $x_2$ and $x_4$, and an arbitrary solution is given by plugging arbitrary values in for these and solving for the pivot variables $x_1$, $x_3$, $x_5$. To get a basis, you plug in $(x_2 = 1$, $x_4 = 0)$ and $(x_2 = 0$, $x_4 = 1)$ to obtain

$$\begin{pmatrix}
-2 \\
1 \\
0 \\
0 \\
\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}
-3 \\
0 \\
1 \\
0 \\
\end{pmatrix}.$$  

More generally, the $i^{th}$ basis element is given by setting the $i^{th}$ free variable to 1 and the rest to zero.

II.B.5. **Example.** Suppose $A$ is an invertible $m \times m$ matrix. I claim that its columns $\{\mathbf{c}_i\}_{i=1}^m$ are a basis for $\mathbb{R}^m$.

To establish this, check that:

- $\{\mathbf{c}_i\}$ are linearly independent: $\gamma_1 \mathbf{c}_1 + \ldots + \gamma_m \mathbf{c}_m = 0 \implies A\mathbf{\tilde{v}} = 0 \implies \mathbf{\tilde{v}} = 0$ (since $A$ is invertible).
- $\{\mathbf{c}_i\}$ span $\mathbb{R}^m$: show every $\mathbf{\tilde{y}} \in \mathbb{R}^m$ can be written $\mathbf{\tilde{y}} = \gamma_1 \mathbf{c}_1 + \ldots + \gamma_m \mathbf{c}_m (= A\mathbf{\tilde{v}})$. This is easy: just put $\mathbf{\tilde{v}} = A^{-1}\mathbf{\tilde{y}}$.  

We say that $V$ is **finitely generated** if some finite subset $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ spans $V$. Here is the precise reason why dimension makes sense:

II.B.6. **Proposition.** If $V$ is finitely generated then (i) it has a (finite) basis, and (ii) any 2 bases of $V$ have the same number of elements.
II.B.7. DEFINITION. We define the dimension \( \dim_F(V) \) of \( V \) over \( F \) to be the number of elements in (ii).\(^5\)

We get (i) by going through the generating subset \( \{\vec{v}_1, \ldots, \vec{v}_m\} \) and throwing out any \( \vec{v}_i \) that is a linear combination of the earlier vectors. What remains is a basis.\(^6\)

In order to show (ii) we first prove the following

II.B.8. LEMMA. If \( \vec{v}_1, \ldots, \vec{v}_m \) span \( V \) then any linearly independent collection in \( V \) has no more than \( m \) vectors.

PROOF. Consider any list \( \vec{u}_1, \ldots, \vec{u}_n, \ n > m \) (of vectors in \( V \)). Since \( \vec{v}_1, \ldots, \vec{v}_m \) span \( V \), each \( \vec{u}_j \) is a linear combination of the \( \vec{v}_i \):

\[
\vec{u}_j = \sum_{i=1}^{m} A_{ij} \vec{v}_i, \quad \text{for } A \in \mathbb{R}^{m \times n}.
\]

Since \( m < n \), there is a nontrivial solution to \( A\vec{x} = 0 \), some nonzero \( \vec{x} \in \mathbb{R}^n \). That is,

\[
\sum_{j=1}^{n} A_{ij} x_j = 0 \quad \text{for each } i,
\]

and so

\[
\sum_{j=1}^{n} x_j \vec{u}_j = \sum_{j=1}^{n} x_j \left( \sum_{i=1}^{m} A_{ij} \vec{v}_i \right) = \sum_{i=1}^{m} \left( \sum_{j=1}^{n} x_j A_{ij} \right) \vec{v}_i = \sum_{i=1}^{m} 0 \vec{v}_i = 0.
\]

We conclude that \( \{\vec{u}_j\}_{j=1}^{n} \) is not independent! \( \square \)

Now let \( \{\vec{v}_1, \ldots, \vec{v}_m\} \) and \( \{\vec{w}_1, \ldots, \vec{w}_n\} \) be two bases for \( V \). The \( \{\vec{v}_i\} \) span \( V \) and the \( \{\vec{w}_j\} \) are independent, so the Lemma implies that \( n \leq m \). The same argument with \( \{\vec{v}_i\} \) and \( \{\vec{w}_j\} \) swapped yields \( n \geq m \), and so we’ve proved Proposition II.B.6. We can rephrase the Lemma as follows: if \( \dim(V) = m \) then (a) less than \( m \) vectors cannot span \( V \), while (b) more than \( m \) vectors are linearly dependent.

---

\(^5\)The \( F \) is typically omitted when there is no ambiguity. That said, in the first example we had \( \dim_{\mathbb{C}}(\mathbb{C}^n) = n \) while \( \dim_{\mathbb{R}}(\mathbb{C}^n) = 2n \).

\(^6\)Any linear dependence relation \( \sum a_i \vec{u}_i = 0 \) on a list of vectors has a largest \( j \) for which \( a_j \neq 0 \). Dividing the relation by \( a_j \) then presents \( \vec{u}_j \) as a linear combination of \( \vec{u}_1, \ldots, \vec{u}_{j-1} \). So if no vector in a list is a linear combination of the vectors preceding it, the vectors are independent.
II. VECTOR SPACES

A Geometric Application. Given 5 points $A, B, C, D, E$ “in general position” in the plane, there is a unique conic passing through them. Here’s why: the equation of a conic $Q$ has the form

$$0 = f_Q(x, y) = \sum_{i=1}^{6} a_i \rho_i^2(x, y) = a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 + a_6 y^2,$$

and the statement that it passes through $A$ is

$$0 = f_Q(x_A, y_A) = \sum a_i \rho_i^2(x_A, y_A).$$

Continuing this for $B$ thru $E$ we get 5 equations in 6 unknowns $a_i$, a system which “generically”\(^7\) has a “line” of solutions: if $(a_1, \ldots, a_6)$ solves the system then so does $(c \cdot a_1, \ldots, c \cdot a_6)$. But this just gives multiples of the same $f_Q$, which won’t change the shape of the curve, and so there is really only one nontrivial solution: through five general points in the plane there exists a unique conic $Q$.

Similarly, given 8 points $A, B, C, D, E, P_1, P_2, P_3$ what can we say about the “vector space” of cubics passing through them (i.e. of cubic polynomials in $(x, y)$ vanishing at all 8 points)? First of all, the vector space of all cubics

$$\sum_{i=1}^{10} a_i \rho_i^3(x, y) = a_1 + a_2 x + a_3 y + \ldots + a_9 x^3 + a_{10} y^3$$

should be thought of in terms of the coefficient vector $\vec{a}$. The 8 constraints give a linear system

$$\begin{align*}
0 &= \sum_{i=1}^{10} a_i \rho_i^3(x_A, y_A) \\
\vdots \\
0 &= \sum_{i=1}^{10} a_i \rho_i^3(x_{P_3}, y_{P_3})
\end{align*}$$

whose space of solutions “generically” is two-dimensional — there are 8 leading 1s in $\text{ref}$ of the $8 \times 10$ $[\rho]$ matrix,\(^8\) and therefore 2 parameters to choose freely.

---

\(^7\)More precisely, “generically” means here that the matrix $[\rho]$ (with rows indexed by $A$ thru $E$, columns by $i = 1$ thru $6$) is of maximal rank $(= 5)$. One can show that this happens when no four of the five points are collinear.

\(^8\)Once again, “generically” means that this matrix has maximal rank.
So start with $A$, $B$, $C$, $D$, $E$ in “general position”; we’d like to construct points on the (unique) conic through them, using a straightedge. Begin by drawing the lines $AB$, $DE$, $CD$, and $BC$; label $AB \cap DE$ by $P_1$. ($Q$ is drawn in a lighter color because we don’t know what it looks like yet.)

Now draw (almost) any line $\ell$ through $A$ — the choice is free except that $\ell$ shouldn’t go through $B$, $C$, $D$, or $E$. All the rest of the construction depends on this choice; in particular the final point $q$ depends on $\ell$, and so by varying $\ell$ we can vary $q$ and in this way (we hope) construct enough points on $Q$ to get an idea of what it looks like. Label $\ell \cap CD =: P_2$; then draw $P_1P_2$ and label $P_1P_2 \cap BC =: P_3$; finally set $q := EP_3 \cap \ell$. We must prove $q \in Q$. Adding dotted lines to our diagram for the lines depending on $\ell$, 


Now consider the following three cubic equations:
\begin{align*}
  f_1(x, y) &:= f_Q(x, y) \cdot f_{P_1P_2}(x, y) = 0, \\
  f_2(x, y) &:= f_{AB}(x, y) \cdot f_{CD}(x, y) \cdot f_{EP_3}(x, y) = 0, \text{ and} \\
  f_3(x, y) &:= f_{\ell}(x, y) \cdot f_{BC}(x, y) \cdot f_{DE}(x, y) = 0
\end{align*}
(where \( f_{AB}(x, y) = 0 \) is the equation of the line \( AB \), and so on). The graphical solutions of these equations are (respectively)
\[ Q \cup P_1P_2, \ AB \cup CD \cup EP_3, \ \ell \cup BC \cup DE. \]
All 3 of these unions contain the 8 points \( A, B, C, D, E, P_1, P_2, P_3 \); that is, the cubic polynomials \( f_1, f_2, \) and \( f_3 \) each vanish at all 8 points. Above we explained that the (vector) space of such cubic polynomials was \( 10 - 8 = 2 \) dimensional;\(^9\) it follows that 3 “vectors” \( f_1, f_2, f_3 \) cannot be linearly independent and there is a nontrivial relation
\[ \alpha f_1(x, y) + \beta f_2(x, y) + \gamma f_3(x, y) = 0. \]
Now we extract the information we want (\( q \in Q \)) from this relation.

First of all \( \alpha \) cannot be zero. Otherwise \( \beta f_2 = -\gamma f_3 \) are multiples of each other, which means that \( AB \cup CD \cup EP_3 = \ell \cup BC \cup DE. \) That would imply \( \ell \) was either \( AB, CD, \) or \( EP_3; \) and we said \( \ell \) cannot pass through \( B, C, D, \) or \( E \) so this gives a contradiction.

So we may divide by \( \alpha \) \( \implies \) \( f_1(x, y) = -\frac{\beta}{\alpha} f_2(x, y) - \frac{\gamma}{\alpha} f_3(x, y). \) Now since \( q \) is contained in both \( EP_3 \) and \( \ell, f_2(x, y) \) and \( f_3(x, y) \) both vanish at \( q = (x_q, y_q). \) (This is because \( f_2 \) contains a factor of \( f_{EP} \) and \( f_3 \) contains a factor of \( f_\ell. \)) According to the relation, therefore also \( f_1 = 0 \) at \( q; \) this means that either \( f_Q \) or \( f_{P_1P_2} \) vanishes at \( q, \) and so \( q \in Q \text{ or } P_1P_2. \) Suppose the latter: if \( q \in P_1P_2 \) then the three lines \( P_1P_2, EP_3, \ell \) are the same, and so \( E \in \ell. \) But once again, we disallowed such a choice of \( \ell \) at the outset. So \( q \in Q, \) and we’re done.

Another way of thinking of this result is that intercepts of opposite edges of a hexagon inscribed in a conic lie on line. This is a theorem proved by Pascal in 1639, and known as his “mystic hexagon”!

\(^9\) It is a (highly nontrivial) fact that the 8 points are “generic” if the first 5 are.
Exercises

(1) Let $V$ be a vector space over a subfield $F$ of the complex numbers. Suppose $\vec{u}, \vec{v}, \vec{w}$ are linearly independent vectors in $V$.

(a) Prove that $\vec{u} + \vec{v}, \vec{v} + \vec{w}$, and $\vec{w} + \vec{u}$ are independent.

(b) Explain why this implies that if $\vec{u}, \vec{v}, \vec{w}$ are a basis, then so are $\vec{u} + \vec{v}, \vec{v} + \vec{w}$, and $\vec{w} + \vec{u}$.

(c) What happens if we take $F$ to be $\mathbb{Z}/2\mathbb{Z}$ instead?

(2) Let $V$ be the set of all $2 \times 2$ matrices $A$ with complex entries which satisfy $A_{11} + A_{22} = 0$.

(a) Show that $V$ is a vector space over the field of real numbers, with the usual operations of matrix addition and multiplication of a matrix by a scalar.

(b) Find a basis for this vector space.

(c) Let $W$ be the set of all matrices $A$ in $V$ such that $A_{21} = -\overline{A_{12}}$ (the bar denotes complex conjugation). Prove that $W$ is a subspace of $V$ and find a basis for $W$.

(3) Prove that the space of all $m \times n$ matrices over the field $F$ has dimension $mn$, by exhibiting a basis for this space.

(4) Let $P_3(x)$ denote the (real) vector space of polynomials (with real coefficients) of degree $\leq 3$ in $x$. Do $2x^2 + 2x + 1$, $x^3 + 2x^2 + 2x + 2$, $x^3 - 4x^2 - 3x - 1$, and $x^3 - 2x^2 - 3x - 4$ constitute a basis of $P_3(x)$? [Hint: you may easily identify $P_3(x)$ with $\mathbb{R}^4$ (how?).]

(5) Given

$$A = \begin{pmatrix} 2 & -3 & -7 & 5 & 2 & -2 \\ 1 & -2 & -4 & 3 & 1 & -2 \\ 0 & 2 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{pmatrix},$$

let $V$ be the vector subspace of $\mathbb{R}^6$ consisting of solutions to $A\vec{x} = 0$. Find a basis of $V$. 