Natural Operators for Linear Algebra

... JUST SCRATCHING AN ITCH.

Heinrich Hartmann
heinrich@heinrichhartmann.com

February 10, 2021

Abstract

In this text we are going to visit basic concepts of linear algebra from the view of an algebraic geometer. Vectors are treated as functions on finite sets which are naturally elements in algebras and modules. We develop a simple formalism of push-forward and pull-back with projection formulas and base-change, duality theorems and adjunctions.

By starting with the infinite case first, we are lead to clear conceptual distinctions between pull-back and push-forward, that are easily overlooked in the finite case.

Matrix calculus is introduced as "kernel-convolutions" on product spaces. Formulas for pull-back and push-forward are derived in terms of matrix product with row/column vectors.

Finally we introduce a syntactic convenience of "covariant composition" that allows us to arrive at perfectly natural formulas for row-vector and push-forward composition, that avoid any mental gymnastics of transposing indices, functions or arguments.

1 Sums and Products

Definition 1.1. Let k be a field and I be a set (possibly infinite). We consider two naturally associated k-vector spaces $S_k(I)$ and $P_k(I)$ defined as follows:

$$P_k(I) = \prod_{i\in I} k = Map(I,k) = \{a : I \to k\}$$

$$S_k(I) = \bigoplus_{i\in I} k = Map_f(I,k) = \{\, a : I \to k \mid a(i) = 0 \text{ almost everywhere } \}$$

Here "almost everywhere" means that there is a finite set $J \subset I$ where the property does not hold. In this case $S_k(I)$ contains functions $I \to k$, that are zero outside of a finite set.

Proposition 1.2 (Basis)**.** For $i \in I$ we have linear maps between these k-vector spaces:

$$e^i : S_k(I) \subset P_k(I) \to k,\ a \mapsto a(i) \qquad e_i : k \to S_k(I) \subset P_k(I),\ 1 \mapsto (j \mapsto \delta_i(j))$$

The elements $e_i := e_i(1)$ are a basis of $S_k(I)$ as k-vector space.

The composition $e_i \circ e^i$ is an idempotent endomorphism of S_k/P_k. The composition $e^i \circ e_i$ is the identity on k. All other compositions are zero.

Proposition 1.3 (Naturality)**.** Let $f : I \to J$ be a map between sets. f induces maps between k-vectors spaces:

$$f_* : S_k(I) \longrightarrow S_k(J),\ a \mapsto (j \mapsto \sum\nolimits_{i:f(i)=j} a(i)) \quad \text{and} \quad f^* : P_k(J) \longrightarrow P_k(I),\ b \mapsto b \circ f$$

For the coordinate maps e_i, e^i this translates into

$$f_* \circ e_i = e_{f(i)}, \quad \text{and} \quad e^i \circ f^* = e^{f(i)}.$$

If $g : J \to K$ is another map, we have:

$$(g \circ f)^* = f^* \circ g^* : P_k(K) \to P_k(I) \quad \text{and} \quad (g \circ f)_* = g_* \circ f_* : S_k(I) \to S_k(K).$$

Proof. For $c \in P_k(I)$, we have $(g \circ f)^*(c) = c \circ g \circ f = f^*(g^*(c))$. For $i \in I$, we have $(g \circ f)_*(e_i) = e_{g(f(i))} = g_*(f_*(e_i))$, hence both maps agree on a basis of $S(I)_k$. □

Example 1.4. Let $I \to \{*\}$ be the projection to a point. Then $f_* : S_k(I) \to k, a \mapsto \sum_i a_i$, is called the *trace map*. The element $\mathbb{1} = f^*(1) \in P_k(I)$ is called *unit*.

Example 1.5 (Bijections)**.** Let $\sigma : I \to J$ be a bijection, then $\sigma_* e_i = e_{\sigma(i)}$ and $\sigma^* e_i = e_{\sigma^{-1}(i)}$. Moreover, $e^i \sigma_* = e^{\sigma^{-1}(i)}$ and $e^i \sigma^* = e^{\sigma(i)}$.

Proposition 1.6 (Adjunction)**.** Let $(\_)^\# : k - Vect \to Set$ be the forgetful functor, that maps a k-vector space to it's underlying set. Then there is a natural isomorphism:

$$Hom_k(S_k(I), V) \cong Map(I, V^\#)$$

In other words S_k is a right-adjoint to $(\_)^\#$. The units/co-units are given by:

$$\varepsilon : S_k(V^\#) \longrightarrow V,\ e_v \mapsto v \quad \text{and} \quad \eta : I \longrightarrow S_k(I)^\#, i \mapsto e_i.$$

Proposition 1.7 (Duality)**.** The k-bilinear pairing

$$(\_,\_) : P_k(I) \times S_k(I) \longrightarrow k, \quad a, b \mapsto (a, b) := tr(a \cdot b) = \sum_i a(i)b(i)$$

is non-degenerate and induces an isomorphism $P_k(I) \to S_k(I)^* := Hom_k(S_k(I), k)$. The dual space of $P_k(I)$ is *not* isomorphic to $S_k(I)$, if I is infinite.

Proposition 1.8 (Linear Adjunction)**.** If $f : I \to J$ is a map, then f_*, f^* are adjoint to each other for the trace pairing:

$$(f^*a, b) = (a, f_*b) \quad \text{for} \quad a \in P_k(J), b \in S_k(I).$$

Proposition 1.9 (Projection Formula)**.** For $f : I \to J$, the projection formula holds:

$$f_*(f^*(a) \cdot b) = a \cdot f_*(b) \quad \text{for} \quad a \in P_k(J), a \in S_k(I).$$

Proof. We have $f_*(f^*(a) \cdot b)(j) = \sum_{i:f(i)=j}(a(f(i))b(i)) = a(j) \sum_{i:f(i)=j}(b(i)) = a \cdot f_*(b)$. □

Proposition 1.10 (Sums)**.** Let $I = I_1 \uplus I_2$ a partition of I. Denote the inclusions by $\iota_i : I_i \to I$. Then

$$\iota_1^* \oplus \iota_2^* : P(I) \longrightarrow P(I_1) \oplus P(I_2) \qquad \iota_{1*} \oplus \iota_{2*} : S(I_1) \oplus S(I_2) \longrightarrow S(I)$$

is an isomorphism.

Proposition 1.11 (Finite maps)**.** A map $f : I \to J$ is called, *finite* if the fibers $f^{-1}\{j\}$ are finite for all j. In this case we have the following identities:

$$e^j \circ f_* = \sum_{i:f(i)=j} e^i, \quad \text{and} \quad f^* \circ e_j = \sum_{i:f(i)=j} e_i.$$

We can extend the definition of f_* from $S(I)$to $P(I)$ in the finite case

$$f_\bullet(a) : P(I) \to P(J), \quad a \mapsto (j \mapsto \sum_{i:f(i)=j} a(j)),$$

so $f_\bullet(a) = f_*(a)$ for $a \in S(I)$.

Similarly, for f^* we have $f^*(S(J)) \subset S(I)$, so that we get an induced map $S(J) \to S(I)$, that we denote by the same symbol f^*.

The projection formula extends to the finite case as

$$f_\bullet(f^*(a) \cdot b) = a \cdot f_\bullet(b), \quad \text{for} \quad a \in P(J), b \in P(I).$$

Example 1.12. For an inclusion $\iota : J \subset I$, we have $\iota_\bullet(\mathbb{1}) = \mathbb{1}_J$, where $\mathbb{1}_J(i) = 1$ if $i \in J$, and 0 otherwise.

Example 1.13. Let $f : I \to J$ be a finite map, then $f_\bullet f^*(b) = f_\bullet(\mathbb{1}) \cdot b$ by projection formula. Here $f_\bullet(\mathbb{1})(j) = \#f^{-1}\{j\}$ counts the cardinality of the fibers.

The other composition can be computed as $f^* f_\bullet(a)(i) = \sum_{i':f(i')=f(i)} a(i')$.

2 Commutative Algebra

Proposition 2.1. The k-vector space $P(I)$ is a k-algebra with point wise multiplication $(a \cdot b)(i) = a(i) \cdot b(i)$) and unit element $\mathbb{1}$. The k-vector space $S(I)$ is an $P(I)$-sub-module of $P(I)$.

Proposition 2.2. If $f : I \to J$ is a map, then f^* is a morphism of k-algebras, and f_* is a morphism of $P(J)$-modules, if we regard $S(I)$ as a $P(J)$ module via f^*.

Proposition 2.3. The idempotent element in $P(I)$ are exactly the functions $\{ e_i \mid i \in I \}$.

Proposition 2.4. The k-algebra morphisms $P(I) \to K$ are exactly the function $\left\{ e^i \mid i \in I \right\}$.

Proposition 2.5. If $A : P(J) \to P(I)$ is a morphism of k-algebras, i.e. $A(x \cdot y) = A(x) \cdot A(y)$, $A(\mathbb{1}) = \mathbb{1}$, then there is a unique $f : I \to J$, with $A = f^*$.

Proposition 2.6. The datum of a $P(I)$-module M, is equivalent to giving a vectors space V with a direct sum decomposition $V = \bigoplus_{i \in I} V_i$.

Proposition 2.7 (Products)**.** If $K = I_1 \times_J I_2$ is a fiber product with structure maps $\pi_i : K \to I_i$, $\sigma_i : I_i \to J$, i.e. $K = \{ (i_1, i_j) \mid \sigma_1(i_1) = \sigma_2(i_2) \} \subset I_1 \times I_2$, then the bilinear map

$$\phi : P(I_1) \times P(I_2) \longrightarrow P(K) \quad a, b \mapsto \pi_1^*(a) \cdot \pi_2^*(b) \tag{1}$$

induces an morphism $P(I_1) \otimes_{P(J)} P(I_2) \to P(K)$ of k-algebras. In case I_1, I_2 are finite, ϕ is an isomorphism.

Proof. Let $a \in P(I), b \in P(J), c \in P(K)$, then

$$\phi(c \cdot a, b) = \pi_1^*(\sigma_1^*(c) \cdot a) \cdot \pi_2^*(b) = (\pi_1 \circ \sigma_1)^*(c) \pi_1^*(a) \pi_2^*(b) = (\pi_2 \circ \sigma_2)^*(c) \pi_1^*(a) \pi_2^*(b) = \phi(a, c \cdot b).$$

So we see that the map (1) is indeed $P(J)$-linear, and induces a linear map $P(I_1) \otimes_{P(J)} P(I_2) \to P(K)$.

Now assume that I_1, I_2 are finite. In this case we can set $\psi(a) = \sum_{(i,j) \in K} a(i,j) e_i \otimes e_j$.

$$\phi(\psi(c)) = \sum_{(i,j) \in K} c(i,j) \pi_1^*(e_i) \cdot \pi_2^*(e_i) = \sum_{(i,j) \in K} c(i,j) e_{(i,j)} = c$$

and

$$\psi(\phi(a,b)) = \sum_{(i,j) \in K} (\pi_1^*(a) \cdot \pi_2^*(b))(i,j) e_i \otimes e_j = \sum_{(i,j) \in K} a(i) b(j) e_i \otimes e_j = \sum_{i \in I_1, j \in I_2} a(i) b(j) e_i \otimes e_j = a \otimes b$$

Where we have used, that $e_i \otimes e_j = 0$ if $(i,j) \notin K$, i.e. $\sigma_1(i) \neq \sigma_2(j)$. ☐

Example 2.8. If $I = \mathbb{N}$, then $\phi : P(I) \otimes_k P(I) \to P(I \times I)$ is not surjective.

Proof. Consider $\Delta \in P(I \times I), \Delta(i,j) = \delta_{i,j}$, and assume $\Delta = \phi(\sum_{\nu=1}^n a_\nu \otimes b_\nu)$. We get have $e_i \otimes e_j \cdot \Delta = \sum_{\nu=1}^n a_\nu(i) b_\nu(j) \cdot e_i \otimes e_j = \delta_{i,j}$. Hence $\sum_{\nu=1}^n a_\nu(i) b_\nu(j) = \delta_{i,j}$ for all $i, j \in \mathbb{N}$. Now consider the map $\alpha : k^n \to P(I)$, defined by $\alpha(x) = (i \mapsto \sum_\nu x_\nu a_\nu(i))$. Then $\alpha(b(i)) = e_i \in P(I)$ for all $i \in \mathbb{N}$. So $Im(\alpha)$ contains $S(I)$, which is impossible since $S(I)$ is infinite dimensional and $dom(\alpha) = k^n$ is not. ☐

Proposition 2.9 (Finite base change)**.** If $K = I_1 \times_J I_2$ is a fiber product with structure maps $\pi_i : K \to I_i$, $\sigma_i : I_i \to J$.

$$\begin{array}{ccc} K & \xrightarrow{\pi_1} & I_1 \\ \big\downarrow{\scriptstyle \pi_2} & & \big\downarrow{\scriptstyle \sigma_1} \\ I_2 & \xrightarrow{\sigma_2} & J \end{array}$$

If $\sigma_1 : I_1 \to J$ is finite, then the same holds for $\pi_2 : K \to I_2$, and in this case

$$\pi_{1*} \pi_2^* = \sigma_1^* \sigma_{2*} : S(I_2) \longrightarrow S(I_1), \quad \text{as well as} \quad \pi_{2\bullet} \pi_1^* = \sigma_2^* \sigma_{1\bullet} : P(I_1) \longrightarrow P(I_2).$$

Proof. If σ_2 is finite, then $i \in I_2$ we have $\pi_2^{-1}\{i\} = \{(i_1, i) | \sigma_1(i_1) = \sigma_2(i)\} \sim \sigma_1^{-1}\{j\}$, with $j = \sigma_2(i)$. Hence π_2 is finite as well.

For $a \in P(I_1)$ we have $\pi_{2\bullet} \pi_1^*(a)(i_2) = \sum_{i_1 : (i_1, i_2) \in K} a(i_1)$, and $\sigma_2^* \sigma_{1\bullet}(a)(i_2) = \sum_{i_1 : \sigma_1(i_1) = \sigma_2(i_2)} a(i_1)$. But the sum conditions are equivalent by the definition of K. The proof for the second identity is near identical. ☐

3 Finite Sets

Definition 3.1. Let I be a finite set, then we have

$$k[I] := S_k(I) = P_k(I).$$

This is a k-vector space with basis $e_i \in k[I]$ and dual basis $e^i : k[I] \to k$. It's a k-algebra with point-wise multiplication and unit $\mathbb{1}$. It comes with the non-degenerate trace-paring $(\_,\_)$ and a trace-map $tr : k[I] \to k, a \mapsto \sum_i a(i)$.

For each map $f : I \to J$, between finite sets I, J we get two adjoint morphisms

$$f^* : k[J] \to k[I], b \mapsto b \circ f \qquad f_* : k[I] \to k[J], e_i \mapsto e_{f(i)}.$$

Definition 3.2 (Products)**.** For finite sets $I_1, \dots, \times I_k$, we set

$$k[I_1, \dots, I_k] := k[I_1 \times \cdots \times I_k].$$

For natural numbers $n_1, \dots, n_k$, we set

$$k[n_1, \dots, n_k] := k[[n_1], \dots, [n_k]]$$

where $[n] = \{1, \dots, n\}$.

Example 3.3. The sets $k[n_1, \dots, n_k]$ are abundant in computational mathematics.

1. Elements in $k[n]$ are often called vectors.
2. Elements in $k[n_1, n_2]$ are often called matrices.
3. Elements in $k[n_1, \dots, n_k]$ are sometimes called tensors.

4 Matrix Calculus

Definition 4.1 (Matrices)**.** Consider the product $I \times J$ with projections π_1, π_2. An element $A \in k[I, J]$ induces two k-linear morphisms:

$$A_* : k[I] \longrightarrow k[J], a \mapsto \pi_{2*}(\pi_1^*(a) \cdot A), \text{ so } (A_*a)(j) = \sum\nolimits_{i \in I} a(i)A(i,j)$$
$$A^* : k[J] \longrightarrow k[I], b \mapsto \pi_{1*}(A \cdot \pi_2^*(b)), \text{ so } (A^*b)(i) = \sum\nolimits_{j \in J} A(i,j)b(j).$$

For the basis vectors we have:

$$A_*e_i = \sum\nolimits_{j \in J} A(i,j)e_j, \qquad A^*e_j = \sum\nolimits_{i \in I} A(i,j)e_i,$$

in other words

$$A_* = \sum\nolimits_{i,j} A(i,j)e_j \circ e^i, \qquad A^* = \sum\nolimits_{i,j} A(i,j)e_i \circ e^j.$$

Definition 4.2 (Matrix composition)**.** If $A \in k[I, J], B \in k[J, K]$ are two matrices. We define their product as

$$A * B := \pi_{13*}(\pi_{12}^*A \cdot \pi_{23}^*B), \text{ so } (A * B)(i,k) = \sum\nolimits_{j \in J} A(i,j)B(j,k)$$

where $\pi_{12} : I \times J \times K \to I \times J, \dots$ are the projections to the factors.

Proposition 4.3 (Associativity)**.** For three matrices $A \in k[I, J], B \in k[J, K], C \in k[K, L]$ we have

$$(A * B) * C = A * (B * C).$$

Proposition 4.4. For matrices $A \in k[I, J], B \in k[J, K]$ the following identities holds:

$$(A * B)_* = B_* \circ A_* : k[I] \longrightarrow k[K] \qquad (A * B)^* = A^* \circ B^* : k[K] \longrightarrow k[I].$$

Proposition 4.5. (Matrix Adjunction) The morphisms A_*, A^* are adjoint for the trace-pairing:

$$(A_*a, b) = (a, A^*b) \quad \text{for all} \quad a \in k[I], b \in k[J]$$

Proof. Indeed $(A_*a, b) = \sum_{i,j} a(i)A(i,j)b(j) = (a, A^*b)$. □

Proposition 4.6. (Matrix representations) Let $\alpha : k[J] \to k[I]$ be a k-linear map, then there exists a unique $A \in k[I, J]$ so that $A^* = \alpha$.

Proof. Set $A(i, j) = e^i \alpha(e_j)$ then $A^* e_j = \sum_i e_i A(i, j) = \sum_i e_i e^i \alpha(e_j) = \alpha(e_j)$. □

Corollary 4.7. Every linear map $\alpha : k[I] \to k[J]$ has an adjoint.

Definition 4.8. (Rows and Columns) Consider the following trivial bijections:

$$row : I \to \{*\} \times I, i \mapsto (*, i), \qquad col : I \to I \times \{*\}, i \mapsto (i, *).$$

The induced maps row_*, col_* map an element $a \in k[I]$, to a matrix with a single row or column, respectively.

For sets I, J we have the transposition operator

$$t : I \times J \to J \times I, (i, j) \mapsto (j, i).$$

Clearly $t \circ row = col$ and $t \circ col = row$.

Proposition 4.9. For $a \in k[I], A \in k[I, J]$, we have:

$$row_*(A_* a) = row_*(a) * A \in k[*, I], \quad col_*(A^* b) = A * col_*(b) \in k[I, *].$$

Proof. We have $(row_*(a) * A)(*, j) = \sum_i a(i)A(i, j) = A_*(a)(j) = row(A_* a)(*, j)$. □

Definition 4.10. (Covariant Composition) The composition formulas get more consistent when we introduce the covariant composition operator. For general maps $f : X \to Y, g : Y \to Z$, and $x \in X$ we define

$$f \bullet g := g \circ f = (x \mapsto g(f(x))) \quad \text{and} \quad x \bullet f := f(x).$$

The covariant composition is clearly associative. It has the advantage, that the composition order is consistent with the direction of the arrows. While this operator might be regarded as trivial, or point-less, it greatly reduces the mental gymnastics required when translating diagrams to formulas, and thus adds considerable convenience.

For $a \in k[I], A \in k[I, J], B \in k[J, K]$, we find:

$$A_* = \sum_{i,j} A(i, j) e^i \bullet e_j, \quad (A * B)_* = A_* \bullet B_*$$

$$(a \bullet A_*)(j) = \sum_i a(i) A(i, j), \quad row(a \bullet A_*) = row(a) * A$$

Moreover

$$e_i \bullet A_* = \sum_j A(i, j) e_j$$

Remark 4.11. Following common conventions we identify $a \in k[I]$ with $row(a) \in k[*, I]$, and denote $col(a)$ by $t(a) = a^t$. With this notation and $A \in k[I, J], B \in k[J, K], c \in k[K]$ we have

$$A^* \circ B^* \circ c = A * B * c^t \quad \text{and} \quad a \bullet A_* \bullet B_* = a * A * B.$$

So matrix pullback A^* and column vectors are "natural" for when using the usual (contravariant) composition operation "$\circ$". For the covariant composition "$\bullet$" row vectors and push-forward are more natural, as illustrated by the above formula.