Kernel composition

1 Definitions

In this note we will assume that the state space S is finite. We can often handle infinite S in a similar way, using the fact that infinite sums of nonnegative numbers behave just like finite sums, as do infinite absolutely convergent sums.

The properties of transition functions can often be expressed conveniently using matrix notation, and matrix notation is also a good way to organize computations. Nevertheless, the geometric aspect of matrices (as rectangular arrays of numbers) is sometimes an unnecessary distraction, as is the requirement that indices must be integers. For that reason we want to have a more abstract notation that captures the essence of the matrix approach.

Definition 1 (The kernel composition operation) *Let a and b be real-valued functions on $S \times S$. Let $a * b$ denote the real-valued function c on $S \times S$ defined by*

$$c(x,z) = \sum_{y \in S} a(x,y)\, b(y,z)\,. \tag{1}$$

We will refer to a, b and c as kernels on S and say that c is the kernel composition of a and b.

If S contains k elements, then a kernel a on S plays the role of a $k \times k$ *matrix* A indexed by the elements of S, so that if $\alpha_{x,y}$ denotes the entry of A corresponding to row index x and column index y, we have

$$\alpha_{x,y} = a(x,y)\,.$$

The kernel composition operation is the analog of *matrix multiplication*, as we can easily check.

In addition to multiplying two $k \times k$ matrices, we often need to multiply a $k \times k$ matrix and a $k \times 1$ column vector or a $1 \times k$ row vector, and we also need the dot product of vectors, which can be thought of as the product of a $1 \times k$ row vector and a $k \times 1$ column vector. The analog of a vector is a real-valued *function* on S. Hence we extend our definition of kernel composition in the following way.

Definition 2 (Extended kernel composition) *Let a be a real-valued functions on $S \times S$ and let f and g be real-valued functions on S. Let $f * a$ be the real-valued function h on $S \times S$ defined by*

$$h(z) = \sum_{x \in S} f(x)\, a(x, z)\,. \tag{2}$$

Let $a * f$ be the real-valued function v on S defined by

$$v(x) = \sum_{z \in S} a(x, z)\, f(z)\,. \tag{3}$$

Let $f * g$ be the number m defined by

$$m = \sum_{x \in S} f(x)\, g(x)\,. \tag{4}$$

2 Properties

Lemma 1 (Associativity) *Let a, b, c be kernels on S, and let f, g be kernels on S. Then*

$$\begin{aligned} (a * b) * c &= a * (b * c)\,, \\ (f * a) * b &= f * (a * b)\,, \\ (f * a) * g &= f * (a * g)\,. \end{aligned} \tag{5}$$

One proves these associativity properties just as one proves that matrix multiplication is associative, so that the proof depends on exchanging the order of summation in a double sum.

Soon we will consider transition functions on the set S. A transition function is a special case of a kernel. The next lemma concerns a slight generalization of a transition function.

Lemma 2 (Sum-preserving kernels) *Let a be a kernel on S, and let w be a real number. Suppose that for every $x \in S$,*

$$\sum_{z \in S} a(x, z) = w. \tag{6}$$

Then the following statements hold.

(i) $a * 1 = w$. Here when we write 1 we mean the constant function whose value everywhere is equal to 1, and similarly when we write w we mean the constant function whose value is w.

(ii) For any function f on S,

$$\sum_{z \in S} (f * a)(z) = w \sum_{z \in S} f(z)\,. \tag{7}$$

(iii) If $a \geq 0$ everywhere, then for any function f on S,

$$\sum_{z \in S} |(f * a)(z)| \leq w \sum_{z \in S} |f(z)|. \tag{8}$$

This remains true if we replace (6) *by*

$$\sum_{z \in S} a(x,z) \leq w. \tag{9}$$

Proof The first statement is obvious from the definition of $*$, and the second statement follows by exchanging the order of summation in a double sum. To prove the third statement, we write

$$\begin{aligned} \sum_{z \in S} |(f * a)(z)| &= \sum_{z \in S} \left| \sum_{x \in S} f(x)\, a(x,z) \right| \\ &\leq \sum_{z \in S} \sum_{x \in S} |f(x)\, a(x,z)| \\ &= \sum_{x \in S} \sum_{z \in S} |f(x)|\, a(x,z) \\ &= w \sum_{x \in S} |f(x)|. \end{aligned} \tag{10}$$

This finishes the proof of the lemma.

Lemma 3 (The constant kernel) *Let u be any real number. We will also use u to denote the constant kernel whose value everywhere is equal to u.*

For any real-valued function f on S, $f * u$ is the constant function ut, where

$$t = \sum_{z \in S} f(z). \tag{11}$$

The proof follows at once from the definition of kernel composition.

Notice that the kernel u corresponds to the matrix U which has every entry equal to u.

Lemma 4 (Transition functions and sums) *Let p be a transition function on S, and let r be a real number. Then the following statements hold.*

(i) $p * 1 = 1$.

(ii) For any function f on S,

$$\sum_{z \in S} (f * p)(z) = \sum_{z \in S} f(z). \tag{12}$$

(iii) For any function f on S,

$$\sum_{z \in S} |(f * p)(z)| \leq \sum_{z \in S} |f(z)|. \tag{13}$$

This lemma is just the special case of Lemma 2 with $w = 1$.

3 Invariant functions

Lemma 5 *Let p be a transition function on S. Let f be a function on S which is invariant, i.e. such that $f * p = f$. Then $|f|$, f^+ and f^- are also invariant.*

Proof We have

$$\begin{aligned} |f(z)| &= |f * p| \\ &= \left| \sum_{x \in S} f(x) p(x, z) \right| \\ &\leq \sum_{x \in S} |f(x)| p(x, z) \\ &= (|f| * p)(z). \end{aligned} \tag{14}$$

That is, $|f| \leq |f| * p$. But, by Lemma 4,

$$\sum_{z \in S} |f(z)| = \sum_{z \in S} (|f| * p)(z). \tag{15}$$

That is,

$$\sum_{z \in S} ((|f| * p)(z) - |f(z)|) = 0, \tag{16}$$

and hence $(|f| * p)(z) - |f(z)| = 0$ for every z. This proves that $|f|$ is invariant.

Since

$$f^+ = \frac{f + |f|}{2}, \ f^{-1} = \frac{|f| - f}{2}, \tag{17}$$

it follows that f^+ and f^{-1} are also invariant, proving the lemma.

Lemma 6 (Existence of invariant distributions) *Let p be a transition function on S. Then there exists an invariant distribution π for p.*

Proof Let P be the transition matrix corresponding to p. As noted in class, the fact that $p * 1 = 1$ implies that $P - I$ is singular. Since row rank equals column rank, there must be a nonzero row vector v such that $v(P - I) = 0$. That is, there must be a nonzero function f such that $f * p = f$.

By Lemma 5, $|f|$ is a nonzero invariant function. Let $\pi = |f|/u$ where $u = \sum_{z \in S} |f(z)|$. Then π is an invariant distribution.

This proves the lemma.

Lemma 7 (Uniqueness of invariant measures) *Let p be a transition function on S. Let $Z_0, Z_1, \ldots$ be the corresponding Markov chain (with any chosen initial distribution).*

Suppose that p is irreducible. This means that for all points $x, y \in S$, if $P(Z_0 = x) > 0$ then for some n we have $P(Z_n = y) > 0$. By the path probability formula, this is equivalent to saying that for some sequence $z_0, z_1, \ldots, z_n$ with $z_0 = x$ and $z_n = y$ we have

$$p(z_0, z_1) \ldots p(z_{n-1}, z_n) > 0. \tag{18}$$

Then p has a unique *invariant distribution.*

Proof By Lemma 6 there is at least one invariant distribution. We must prove uniqueness. Suppose that π and μ are invariant distributions. Then $\pi - \mu$ is an invariant function.

By Lemma 5, $(\pi - \mu)^+$ and $(\pi - \mu)^{-1}$ are invariant. Let $f = (\pi - \mu)^+$ and let $g = (\pi - \mu)^-$.

If f is not the zero function, then $f(x) > 0$ for some x. For any point $y \in S$, by assumption there is some n such that

$$(f * p^{*n})(y) > 0. \tag{19}$$

But $f * p^{*n} = f$. Hence $f(y) > 0$ for all $y \in S$. Therefore g is the zero function.

But since π and μ are both distributions,

$$\sum_{z \in S} (\pi(z) - \mu(z)) = 1 - 1 = 0. \tag{20}$$

Thus

$$\sum_{z \in S} f(z) = \sum_{z \in S} g(z) = 0 \tag{21}$$

and so $f = 0$, contradiction. Similarly we see that $g = 0$.

This proves the lemma.