Convex cones and the Hilbert metric

Vaughn Climenhaga

March 28, 2013

Having spent some time discussing spectral methods and coupling techniques as tools for studying the statistical properties of dynamical systems, we turn now to a third approach, based on convex cones and the Hilbert metric. This post is based on Will Ott's talk from March 25.

1   Basic definitions

Let V be a vector space over the reals. Ultimately we will be most interested in the case when V is a function space, such as L^1 or BV, but for now we make the definitions in the general context.

Definition 1 *A subset $C \subset V$ is a convex cone (or positive cone) if*

1. *$C \cap (-C) = \emptyset$;*

2. *$\lambda C = C$ for each $\lambda > 0$;*

3. *C is convex; and*

4. *for all $f, g \in C$ and $\alpha \in \mathbb{R}$, we have the following property: if $\alpha_n \to \alpha$ and $g - \alpha_n f \in C$ for every n, then $g - \alpha f \in C \cup \{0\}$.*

The first three conditions are very geometric and in some sense guarantee that C "looks like a cone should look". The last condition is more topological; if V is a topological vector space and $C \cup \{0\}$ is a closed subset of V, then this condition holds, but we stress that the condition itself is actually weaker than this and is phrased without reference to any topology on V.

Example 1 *Let $V = BV([0,1], \mathbb{R})$ be the space of all real-valued functions on the unit interval with bounded variation, and let $C = \{\varphi \in V \mid \varphi \geq 0, \varphi \not\equiv 0\}$. Then C is a convex cone.*

We see immediately from this example that the notion of convex cone is relevant to the sorts of questions we want to ask about invariant measures of a dynamical system, because this set C is exactly the set of density functions that arises when we are searching for an absolutely continuous invariant measure.

This suggests that we will ultimately want to consider the action of some operator $L\colon C \to C$, and in particular may want to find a fixed point of this action (for a suitable operator L). One of the most powerful methods for finding a fixed point is to find a metric in which L acts as a contraction, and this is accomplished by the Hilbert metric, which we now introduce.

Definition 2 *Fix a convex cone $C \subset V$. Given $\varphi, \psi \in C$, let*

$$
\begin{aligned}
\beta(\varphi, \psi) &= \inf\{\mu > 0 \mid \mu\varphi - \psi \in C\}, \\
\alpha(\varphi, \psi) &= \sup\{\lambda > 0 \mid \psi - \lambda\varphi \in C\},
\end{aligned}
\tag{1}
$$

with $\alpha = 0$ and/or $\beta = \infty$ if the corresponding set is empty. The cone distance between φ and ψ is

$$
d_C(\varphi, \psi) = \log\left(\frac{\beta(\varphi, \psi)}{\alpha(\varphi, \psi)}\right).
\tag{2}
$$

The distance d_C is also called the Hilbert (projective) metric.

Several remarks are now in order. First we observe that although V may be infinite-dimensional, the distance $d_C(\varphi, \psi)$ is completely determined in terms of the two-dimensional subspace spanned by φ and ψ, and in particular by the points shown in Figure 1 – in the figure, the lines $0A$ and $0B$ are the boundary of this two-dimensional cross-section of C. The lines $0X$ and $Y\psi$ are parallel, as are the lines $0A$ and ψX; then we have

$$
\alpha = \frac{|\psi Y|}{|0\varphi|} \text{ and } \beta = \frac{|0X|}{|0\varphi|}.
$$

An alternate description of d_C is available in terms of this more geometric description. Let ℓ be the line through φ and ψ, and let A, B be the points
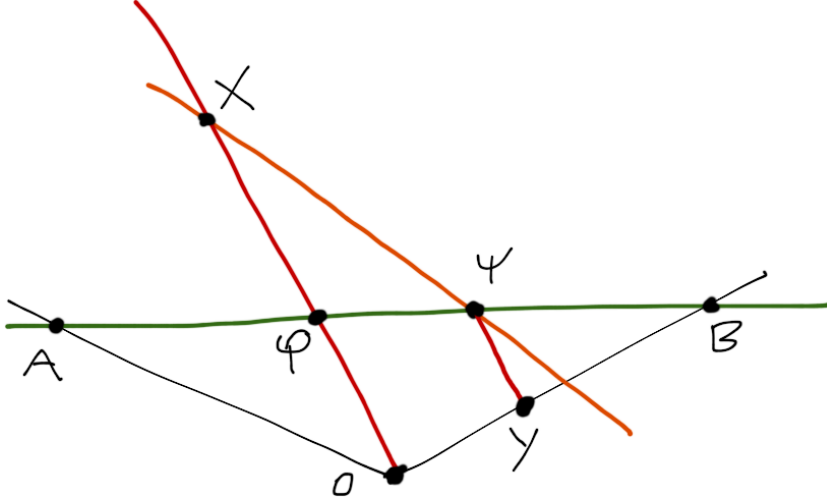
Figure 1: Determining the cone distance between φ and ψ.

where this line intersects the boundary of C. We see from Figure 1 that the triangles $BY\psi$ and $B0\varphi$ are similar, so

$$\alpha = \frac{|\psi Y|}{|0\varphi|} = \frac{|B\psi|}{|B\varphi|}.$$

Furthermore, $\varphi 0 A$ and $\varphi X \psi$ are similar, so

$$\beta = \frac{|0X|}{|0\varphi|} = 1 + \frac{|\varphi X|}{|0\varphi|} = 1 + \frac{|\psi\varphi|}{|A\varphi|} = \frac{|A\psi|}{|A\varphi|}.$$

Thus d_C can be given in terms of the cross-ratio of the points φ, ψ, A, B:

$$\frac{\beta}{\alpha} = \frac{|A\psi|}{|A\varphi|}\frac{|B\varphi|}{|B\psi|} = (\varphi, \psi; A, B).$$

We have

$$d_C(\varphi, \psi) = \log(\varphi, \psi; A, B). \tag{3}$$

Note that it is possible that the line ℓ does not intersect the boundary of C twice; this corresponds to the case when either $\alpha = 0$ or $\beta = \infty$ (or both) in (1), and in this case $d_C(\varphi, \psi) = \infty$.

This situation occurs, for example, when we take $V = BV([0,1], \mathbb{R})$ and C as in the example above, and consider $\varphi, \psi \in C$ with disjoint supports –

3

that is, $\varphi(x)\psi(x) = 0$ for all x. In this case $\alpha = 0$ and $\beta = \infty$ so the cone distance between φ and ψ is infinite.

Because of this phenomenon, d_C is not a true metric. Moreover, we observe that d_C is projective: $d_C(\varphi, \lambda\varphi) = 0$ for every $\lambda > 0$.

An important property of the Hilbert metric is the following theorem, due to Birkhoff, which states that a linear map from one convex cone to another is a contraction whenever its image has finite diameter.

Theorem 3 *Let $C_1 \subset V_1$ and $C_2 \subset V_2$ be convex cones, and let $L\colon V_1 \to V_2$ be a linear map such that $L(C_1) \subset C_2$. (This is a sort of 'positivity' condition.) Let*

$$\Delta = \sup_{\hat{\varphi},\hat{\psi}\in L(C_1)} d_{C_2}(\hat{\varphi}, \hat{\psi}).$$

Then for all $\varphi, \psi \in C_1$, we have

$$d_{C_2}(L\varphi, L\psi) \leq \tanh\left(\frac{\Delta}{4}\right) d_{C_1}(\varphi, \psi), \tag{4}$$

where we use the convention that $\tanh\infty = 1$.

We also want to relate d_C to a more familiar norm. Say that a norm $\|\cdot\|$ on V is *adapted* if the following is true: whenever $\varphi, \psi \in V$ are such that $\varphi - \psi \in C$ and $\varphi + \psi \in C$, we have $\|\psi\| \leq \|\varphi\|$.

Example 2 *On BV, the L^1 norm is adapted, but the BV norm is not.*

The following lemma, due to Liverani, Saussol, and Vaienti, relates the cone metric to an adapted norm.

Lemma 4 *Let $\|\cdot\|$ be an adapted norm on V and $C \subset V$ a convex cone. Then for all $\varphi, \psi \in C$ with $\|\varphi\| = \|\psi\| > 0$, we have*

$$\|\varphi - \psi\| \leq \left(e^{d_C(\varphi,\psi)} - 1\right) \|\varphi\|. \tag{5}$$

Convex cones and the Hilbert metric are well suited to studying nonequilibrium open systems. Consider the following setting. Let X be a Riemannian manifold, λ volume on X, and $\hat{f}_i\colon X \to X$ a diffeomorphism. For $m \in \mathbb{N}$, let $\hat{F}_m = \hat{f}_m \circ \cdots \circ \hat{f}_1$. This is a nonequilibrium closed system. (Nonequilibrium because the map changes at each time step, closed because every point can be iterated arbitrarily many times.)

Now consider sets $H_j \subset X$, which we interpret as a "hole" at time j. The *time-m survivor set* is

$$S_m = X \setminus \bigcup_{i=1}^{m} \hat{F}_i^{-1}(H_i),$$

the set of points that do not fall into a hole before time m. Let $F_m = \hat{F}_m|_{S_m}$. We refer to the pair (F_m, H_m) as a *nonequilibrium open dynamical system*.

We would like an analogue of decay of correlations for such systems. Let φ_0, ψ_0 be two probability density functions on X, and evolve these under (F_m, H_m). We expect that $\|\varphi_t\|_{L^1(\lambda)} < 1$ because there is a positive probability of falling into a hole.

Let $\hat{\mathcal{P}}_j$ be the Perron–Frobenius operator for the closed system $\hat{f}_j$ (with respect to λ). Then to the open system f_j we can associate the operator

$$\mathcal{P}_j(\varphi) = \hat{\mathcal{P}}_j(\varphi)\mathbf{1}_{X \setminus H_j}.$$

Definition 5 *We say that (F_m, H_m) exhibits conditional memory loss in the statistical sense if for all suitably chosen φ_0, ψ_0, we have*

$$\lim_{t \to \infty} \left\| \frac{\varphi_t}{\|\varphi_t\|_{L^1(\lambda)}} - \frac{\psi_t}{\|\psi_t\|_{L^1(\lambda)}} \right\|_{L^1(\lambda)} = 0.$$

The idea of this definition is that before comparing the probabilities, we need to first condition on the event that the trajectory survives. Next time we will investigate this property for piecewise expanding interval maps using the Lasota–Yorke inequality, where the holes H_j are small and vary slowly.