Fraud Detection, Quantum Mechanics, and Complex Systems Lecture 3 Notes for CSSS2010

Gregpry Leibon

June 18, 2010

Disclaimer: Same declaimer as lecture 1 (Complex Systems, Statistical Learning, and Pickles). The material from lecture 1 is freely used.

Okay, okay, the quantum part of the title is something of joke, mathematicians simply can't resist the word quantum! But in all seriousness, I hope to convince you the tools from quantum mechanics play an interesting role in exploring the geometry of a Markov chain. Equally importantly this analogy allows us the opportunity to dwell on a fundamental issue in in the study of complex systems, that the environment is often determined by the objects in the system and visa versa. Here we use the mathematical tools from basic quantum mechanics to discuss and quantify this. Concretely, we view our system as being described by a *state in a Hilbert space*, which on this case is nothing more more than a complex valued function on the chain's states viewed as living in $\mathbb{C}^{|States|}$ which has been endowed with the natural Hermitian inner-product

$$< f, g >_{\omega} = \bar{f}^{tr}[\omega]g.$$

Notice, that the Green's embedding lying at the heart of the previous chapter was conformally invariant, hence there was no dependency on the equilibrium measure and no dependency on this inner product. By contrast, in this section this inner product is of paramount importance.

The notion of context discussed above is encoded via the eigenfunctions and eigenvalues of a Hermitian operator on the state space called the *context operator* denoted C_{ω}. Objects are also determined by an operator, but it worth first considering what an object is. It is a some sort of coherent packet of information; perhaps a motif, a pattern, a scheme, or a concept. We will call such coherent objects *scenarios* to embrace their generality, and once again they are encoded via the eigenfunctions and eigenvalues of a Hermitian operator, called the *scenario operator* and denoted S_{ω}. Now the eventual goal is to see how coupled a scenario is to the context relative to a given state, and this concept is once encoded via the eigenfunctions and eigenvalues of a Hermitian operator which we might call the *anomaly operator*.

Okay, okay, this probably looks a bit nutty, and to make this concrete and introduce the characters relevant to the Markov chain story being told here, will require us to first improve our understanding of chains (recall for us chains are continuous time, ergodic Markov chains with $p_{ii} = 0$.)

1 Conductance and flow

In this section, we see how to view a Markov chain as conductance together with a flow. Recall our toy chain constructed in lecture 1 was derived from a conductance matrix W, and it is best to begin such a discussion with what makes such chains special. To understand why, we first note that we can run the chain backwards in time using the discrete chain

$$P^* = [\pi]^{-1} P^{tr} [\pi],$$

and we define a chain to be *reversible* if $P = P^*$. In exercise 1 we explore some fact about reversibility including the fact that an chain is reversible if and only if it can be derived from a conductance matrix. In other words, there is a symmetric matrix W such that $P = [W\mathbf{1}]^{-1} W$. Given W computing the equilibrium is easy as $\pi = W\mathbf{1}/(\mathbf{1}^{tr} W\mathbf{1})$, while given a reversible chain P and it's equilibrium vector π we have $W = [\pi]P$. The real key is that one can interpret W as the conductance of a electric network, see [2].

A not necessarily reversible chain also has an underlying conductance matrix W as well as a compatible divergence free flow F. *Compatible* means $c_{ij} \leq f_{ij}$, a *flow* means it is skew symmetric, and *divergence free* means $F\mathbf{1} = \mathbf{0}$. Given W and F

$$P = [W\mathbf{1}]^{-1} (W + F),$$

$$\pi = W\mathbf{1} / \left(\mathbf{1}^{tr} W\mathbf{1}\right),$$

and

$$P^* = [W\mathbf{1}]^{-1} (W - F).$$

Conversely given π and P

$$W = [\pi] \frac{P + P^*}{2}$$

and

$$F = [\pi] \frac{P - P^*}{2}.$$

Notice, the Laplacian of a reversible chain satisfies

$$\Delta_\omega^* = [\omega]^{-1} \Delta_\omega^{tr} [\omega],$$

and the star can be interpreted (as it often is) ad the adjoint, in other words

$$\langle f, \Delta_\omega g \rangle = \langle \Delta_\omega^* f, g \rangle.$$

Example: Non-reversible toy For example, we can form the non-reverble chain by using the conductance of our Toy Example and assining a divergecne free flow. The cycles in figure 1 are such flows. We can

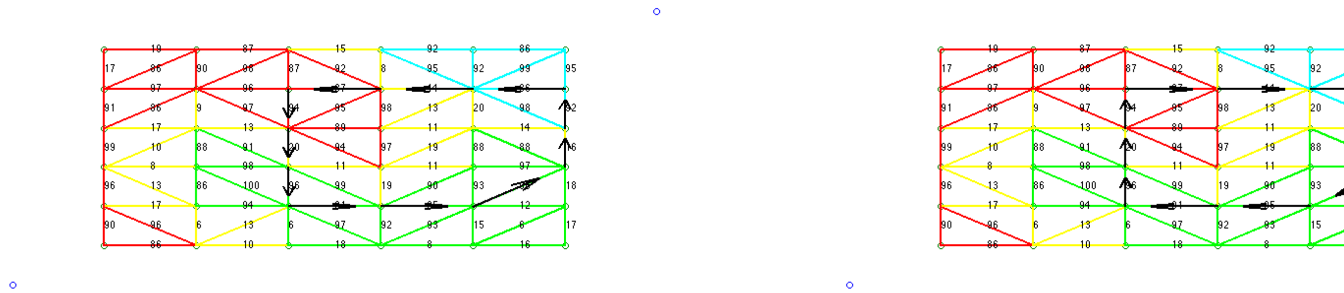

Figure 1: Here we see our toy chain made non-reversible by describing its new P using it's original conductance F and a flow determined by the cycle in the picture. We see its time reversal P^* as well.

assign to the flow an constant weight (divergence free) that is less than or equal to the largest edge weight (compatible).

Exercises:

1. Recall our chains are ergodic with $p_{ii} = 0$ and that we defined reversible as $P = P^*$.

   (a) Prove for every chain there exists a conductance matrix W (unique up to a global constant) and a flow F as defined in the non-reversible toy example such that $P = [W\mathbf{1}]^{-1}(W+F)$, and that $\pi = [W\mathbf{1}]/(\mathbf{1}^{tr}W\mathbf{1})$ and $P^* = [W\mathbf{1}]^{-1}(W-F)$.

   (b) Prove being reversible is equivalent to $F = 0$.

   (c) Prove that reversible is equivalent to for any sequence of sates $a_1,...,a_n$ that the probability of this sequence occurring in equilibrium is the same as the sequence $a_n,...,a_1$.

   (d) Prove that reversible is equivalent to the expected time that it takes to visit the states in $a_1,...,a_n$ in that order is the same as the expected time to visit the states in $a_n,...,a_1$ in that order.

2. Prove reversibility is equivalent to the cross potential satisfying the symmetry $[a,b;c,d] = [c,d;a,b]$.

3. (100$ for first correct solution/counter example sent to me at gleibon/@gmail.com by anyone other than Peter Doyle) Call two chains commute equivalent if for all a and b they share the same expected commute time. Is every chain commute equivalent to a reversible chain?

2 The context operator

Here we define an operator that defines context in our chain. We will choose a rather special context operator deeply related to the clustering of the chain. Or more precise to the rate at which we leave a region R. To define this, we first look at the state vectors that defines a region

$$\chi_R(i) = \begin{cases} \frac{1}{\sqrt{\omega(R)}} & \text{if } i \in R \\ 0 & \text{if } i \in \bar{R} \end{cases}$$

Notice, if R_1 and R_2 are disjoint $\langle \chi_{R_1}, \chi_{R_2} \rangle_\omega = \delta_1^2$, or rather the χ_R are unit vectors with, and χ_{R_1} and χ_{R_2} are orthogonal if R_1 and R_2 are disjoint.

Vagabond Theorem:

$$P(X_{t+dt} \in \bar{R} \mid X_t \in R) \approx ||\chi_R||^2_{Dir}\, dt$$

proof:

$$\begin{aligned}
P(X_{t+dt} \in \bar{R} \mid X_t \in R) &= \frac{P(X_{t+dt} \in \bar{R}, X_t \in R)}{P(X_t \in R)} && \text{(def. conditional prob.)} \\
&= \frac{\sum_{i \in \bar{R}, j \in R} P(X_{t+dt} = i, X_t = j)}{\omega(R)} && \text{(def.)} \\
&= \frac{\sum_{i \in \bar{R}, j \in R} P(X_{t+dt} = i | X_t = j) P(X_t = i)}{\omega(R)} && \text{(def. conditional prob.)} \\
&= \frac{\sum_{i \in \bar{R}, j \in R} (1/\tau_j) p_{ij} dt \omega_i}{\omega(R)} && \text{(instantaneous transition rate)} \\
&= \frac{\sum_{i \in \bar{R}, j \in S} p_{ij} \pi_i}{\omega(R)} dt && (\omega_i = \tau_j \pi_i) \\
&= \textstyle\sum_{i \in \bar{R}, j \in R} \pi_i p_{ij} (\chi_R(i) - \chi_R(j))^2 dt && \text{def.} \\
&= (1/2) \textstyle\sum_{i,j} \pi_i p_{ij} (\chi_R(i) - \chi_R(j))^2 dt && \text{(summand 0 if } i,j \in S \text{ or } i,j \in \bar{R}) \\
&= \langle \chi_R, \Delta_\pi \chi_R \rangle_\pi dt && \text{Green's Identity} \\
&= ||\chi_R||^2_{Dir}\, dt && \text{conformal invariance}
\end{aligned}$$

Q.E.D

Armed with this theorem, we imagine that an understanding of the context should allow one to solve problems akin to finding a partition $\{R_i\}_{i=1}^K$ of our states into regions that minimizes the rate at which we leave the parts. In other words, solve the...

The Vagabond Problem: Finding a partition $\{R_i\}_{i=1}^K$ of our states into regions that minimizes

$$\sum_i P(X_{t+dt} \in \bar{R}_i \mid X_t \in R_i) = \sum_i ||\chi_{R_i}||^2_{Dir}.$$

Now we can relax the Vagabond Problem and search for a ω-orthonormal set $\{f_i\}_{i=1}^L$ such that

$$\sum_i^L ||f_i||^2_{Dir}$$

is minimized. With such $\{f_i\}_{i=1}^L$ we could then find the regions by applying K-Means to the embedding determined by the $\{f_i\}_{i=1}^L$ functions (as with spectral clustering), see exercise 1). We call the embedding via these f_i functions the *Vagabond Embedding*. To actually find the Vagabond Embedding, notice that $< f_i, \Delta_\omega f_i >_\omega = ||f_i||^2_{Dir}$ is being used as a quadradic form, hence each term in the sum remains the same if we view Δ_ω an the Hermitian operator

$$C_\omega = \frac{\Delta_\omega + \Delta_\omega^*}{2}.$$

In other words,

$$< f, \Delta_\omega f >_\omega = < f, C_\omega f >_\omega$$

and we have replaced the minimization problem with minimizing

$$\sum_i \frac{< f_i, C_\omega f_i >_\omega}{< f_i, f_i >_\omega}.$$

C_ω is our *context operator* and being Hermitian allows us to use the spectral theorem and the Raileigh-Rtiz method to identify our Vagabond Embedding as the ω-unit length eigenfunctions of C_ω with the smallest eigenvalues.

Flow Independence Corollary: The vagabond embedding and C_ω depend only on on the Conductance W and the equilibrium measure ω, and are independent of the flow F.

proof: Simply note

$$C_\omega = \frac{\Delta_\omega + \Delta_\omega^*}{2} = [\tau]^{-1}\left(I - \frac{P+P^*}{2}\right) = [\tau]^{-1}\left(I - [\pi]^{-1}W\right)$$

we have:
Q.E.D
Exercises:

1. In the special case, of a reversible chain $C_\omega = \Delta_\omega$ as used in spectral clustering (see **??**ex.**??**). Explain the sense in which we have just generalize Spectral Clustering from section **??** to nonreversible and continuous time Markov Chains.

2. Here we comparing the vagabond embedding and the Green's Embedding.

   (a) Prove, in the reversible case prove that Vagabond Embedding is the PCA of the Green's embedding weighted by the equilibrium measure.

   (b) Hence one can produce embedding of a reversible chain using the Vagabond embedding or via a weighted PCA of the Green's embedding. Do these embeddings agree? If not, produce an example where they disagree.

3. Our choice of the context here is perhaps the simplest context. But any operator will do, though one that is not localized would be most natural. Develop a second context operator.

2.1 Vagabond embedding examples

Toy Example: In figure 2 we see the vagabond embedding our toy chain in the $\tau_i = 1$ case. Clearly the clusters do correspond to the colored components as one would hope. Notice that the embedding is VERY different from the Green's embedding. One reason, is that the outliers have large commute time because they are hard to get to. For the same reason if we keep the equilibrium measure in mind the we can mitigate for this and de-emohasize the outliers. In exercise 2ex.2, this is made entirely explicit where you are ask to show that (in the reversible) case that the vagabond embedding **is** the weighted PCA of the Green's embedding weighted by the equillibrium measure. Notice, we are minimizing the quotient $||f_i||_{Dir} / ||f_i||_{\omega}$, and despite the fact that the numerator $||f_i||_{Dir}$ is conformally invariant, the underlying equilibrium measure is plays a fundamental role there in the denominator. As is always this case with quadratic forms, eigenfunctions, and eigenvalues: *it takes to too tango*.

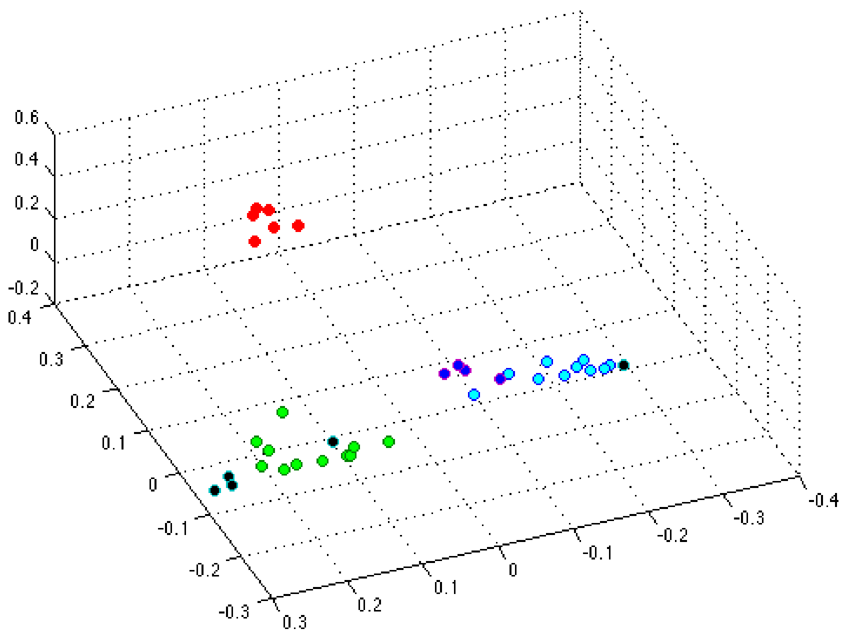

Figure 2: The vagabond embedding of the toy chain with $\tau_i = 1$.

It takes two to Tango: In the previous example we saw the extremely important role of the equillibrium measure with regard to the vagabond embedding. This is telling us that the time we spend at a state, the τ_i, are **crucial** to the geometry. Here we look at a couple of examples with regard to our toy chain. Will use $\tau = \frac{\tau_0}{\sum \tau_0 \pi}$ with the below τ_0. If we slow down the blue corner in the lower left of figure and speed up the magenta triagle in the upper right via $\tau_0(blue) = 5$, $\tau_0(magenta) = 1/5$, and $\tau_0(other) = 1$ then we get a the vagabond embedding on the left of figure 3 where indeed the dark blue vertexed triangle that we are now spending a lot of time at becomes a cluster. If we speed both these regions up via $\tau_0(blue) = 1/5$, $\tau_0(magenta) = 1/5$, and $\tau_0(other) = 1$ then we get a the vagabond embedding on the left of figure 3 where

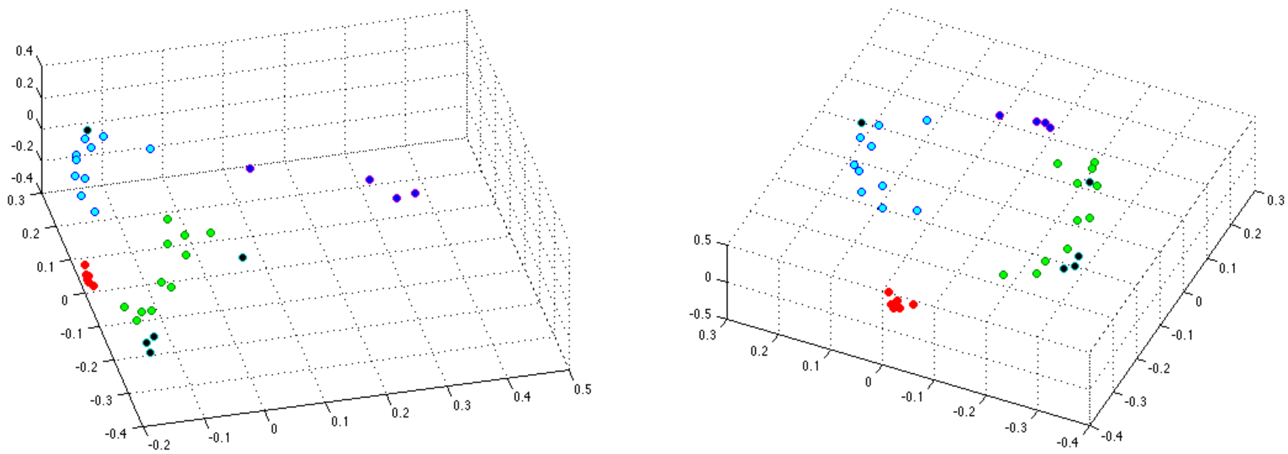

Figure 3: Conformal changes of our toy change.

the blue triangle has migrated to the green region. The up shot is that we now clearly see that the time we wait before moving plays a significant role in understanding the geometry.

Space of Mathematics Example: In figure 4 we see the $\tau = 1$ vagabond embedding of the Space of Mathematics. On the left in figure 4 we use the distance $\sqrt{E(T_{ab})}$ and MDS to embed a chunk of the space of mathematics into R^3. To test the meaning of the embedding, we randomly selected math pages from three rather distinct parts of mathematics using Wiki's *'list-of-computability-and-complexity-topics'* and mark them green, *'list-of-geometric-topology-topics'* and mark them blue, and *'list-of-statistics-articles'* and mark them black (with a touch of red). As with our toy example, directly MDSing Green's embedding will leave you with the hard to reach outliers dominating the visualization (thought the structure is still evident when you look closely). On the right we see the vagabond embedding of the space, and as before it is VERY different from the Green's embedding. As with our toy chain the Vagabond embedding once again emphasize the global clustering properties we'd expect from the context operator.

Exercises:

1. In the Space Of Mathematics Example we assume $\tau = 1$. How can we make a better choice for the time spent at a page? Implement it.

3 The scenario operator

Here we are going explore a scenario in our space. This scenario is going to be 'cycle like behavior', and hence we must ask: what does a directed cycle look like from the point of view of the functions on the chain?

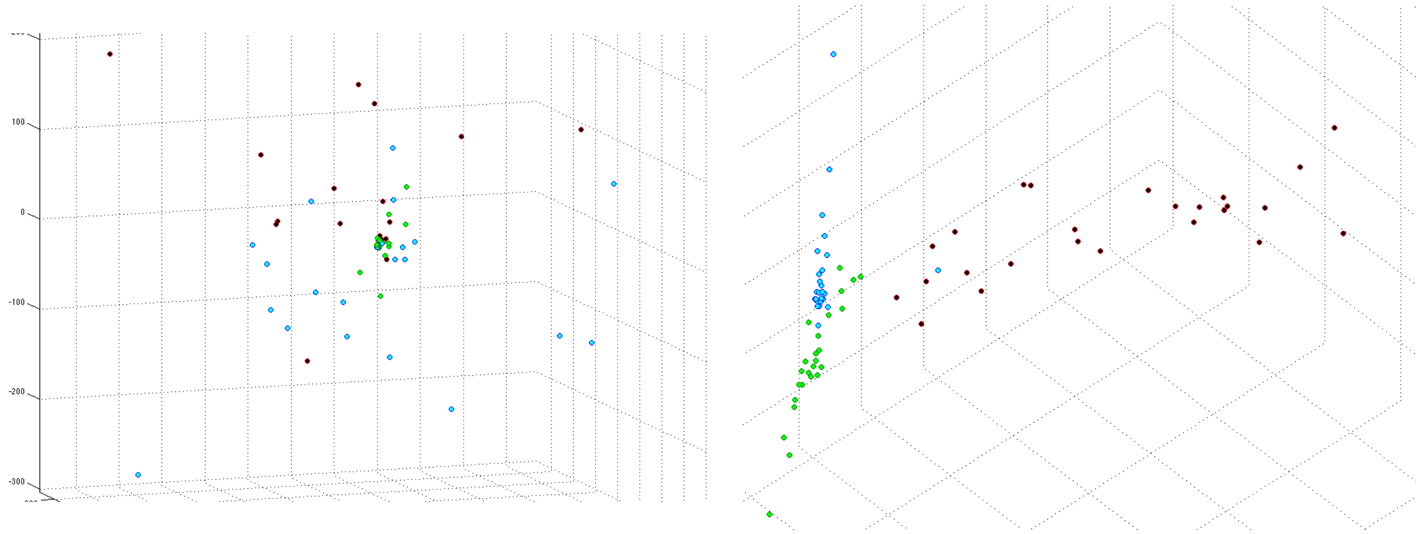

Figure 4: On the left we have the Green's embedding of the Space of Mathematics. On the right we have the Vagabond embedding of the Space of Mathematics with $\tau_i = 1$.

To oreint our selves let's look at a cycle chain with N states where we travel from state $n = 0$ to $n = N-1$. Notice the function $F(n) = e^{i\left(\frac{K2\pi}{N}\right)n}$ satisfies

$$e^{-i\left(\frac{K2\pi}{N}\right)n} P^n f = f.$$

More generally, cycles will support functions where there is a relatively large frequency κ such that

$$e^{-i\kappa t} e^{-t\Delta} f$$

is pointwise relatively stable. One way to find largish κ such the function f and κ pair locally minimize $\left|\left|\frac{d}{dt} e^{-(i\kappa+\Delta)t} f\right|\right|_\omega^2$ at $t = 0$. So we are looking for critical points of

$$F(\kappa, f) = \left|\left|\frac{d}{dt} e^{-(i\kappa+\Delta)t} f\right|\right|_\omega^2\Bigg|_{t=0}.$$

We might call these local minima cycle functions. To get a sense for 'largish', putting a cycle on our toy chain as in figure 1, we can compare the function in the figure 5, which for $\kappa = 0.541126346528966$ has a nearly zero derivative, to a random function (real and imaginary parts random in $[0,1]$ normalized to have norm 1). For a random function we find its optimal κ for minimizing the derivative and after 10000 runs the largest such $\kappa < 10^{-14}$.

We found this function and can find many other with the following lemma:

Cycle Function Detection Lemma: If $S_\omega f = \kappa f$ where $S_\omega = i\frac{\Delta - \Delta^*}{2}$, then $\frac{d}{d\kappa} F = 0$.

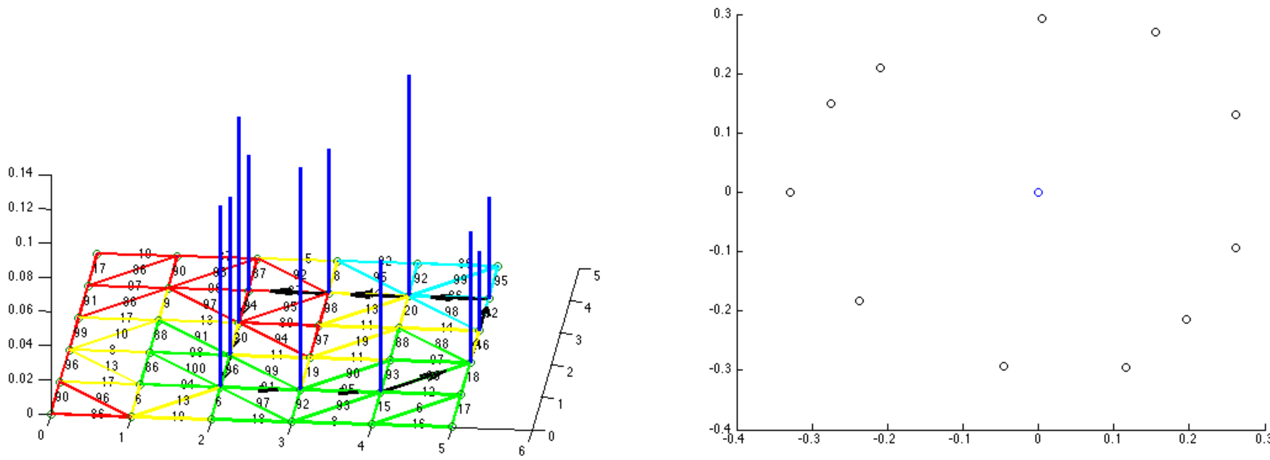

Figure 5: A cycle function on our toy chain with a cycle. On the left is is absolute value, on the right is plot of the values of the eigenfunction.

proof:

$$\begin{aligned}
0 &= \frac{d}{d\kappa}\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}\Bigg|_{t=0} && \text{(definition)}\\
&= \frac{d}{d\kappa}\left|\left|(i\kappa+\Delta)e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}\Bigg|_{t=0} && \text{(matrix derivative)}\\
&= \frac{d}{d\kappa}||(i\kappa+\Delta)f||_{\omega}^{2} && (e^{At}=I \text{ at } t=0\ )\\
&= \frac{d}{d\kappa}\langle(i\kappa+\Delta)f,(i\kappa+\Delta)f\rangle_{\omega} && (\text{def. } <,>_{\omega}\ )\\
&= \langle if,(i\kappa+\Delta)f\rangle_{\omega}+\langle(i\kappa+\Delta)f,if\rangle_{\omega} && \text{(chain rule)}\\
&= \langle f,(\kappa-i\Delta)f\rangle_{\omega}+\langle(\kappa-i\Delta)f,f\rangle_{\omega} && \text{(conjugate linear)}\\
&= \langle f,((\kappa-i\Delta)+(\kappa-i\Delta)^{*})f\rangle_{\omega} && \text{(def. adjoint)}\\
&= \langle f,2(\kappa-S_{\omega})f\rangle_{\omega} && \text{(compute adjoint)}
\end{aligned}$$

Q.E.D

and it is indeed self adjoint

$$\langle f,S_{\omega}g\rangle=\langle S_{\omega}f,g\rangle$$

so the Spectral Theorem assures us that we an ω orthonormal basis of eigenfunctions with real eigenvalues. So the goal becomes to find the larger eigenvalued eigenfunction of the *scenario operator*

$$S_{\omega}=i\frac{\Delta-\Delta^{*}}{2}.$$

In the last section, it was observed that our cluster context operator C_{ω} depends only on the reversible aspects of the chain, in other words via ω and the conductance W. The opposite is true for our cycle scenario operator S_{ω}.

Condctance Independence Corollary: S_ω depends only on the flow F and equilibrium vector ω and is independent of conductance W. Furthermore the chain is reversible if and only if $S_\omega = 0$.

proof:

$$S_\omega = i\frac{\Delta_\omega - \Delta_\omega^*}{2} = i[\tau]^{-1}\left(\frac{P-P^*}{2}\right) = i[\tau]^{-1}[\pi]^{-1}F = i[\omega]^{-1}F$$

Q.E.D

Cycles in the Space of Mathematics: We found cycles in space of mathematics. In figure 6, we see an eigenfunction. Which contains the cycle

$$\begin{aligned}&'hyperbolic_plane',\\ &'isotropic\_quadratic\_form',\\ &'witt\%27s\_theorem',\\ &'witt\_group'\end{aligned}$$

determined by the The co-conformal cycle hunt algorithm. This cycle involves the hyperbolic plane, a key concept in the previous chapter.

3.1 The co-conformal cycle hunt

The eigenfunctions with high absolute value are very localized for our scenario operator. However, the cycles themselves are far from being in one to one correspondence with the eigenfunctions. In figure 7 we see an example of this phenomena where an eigenfunction is a linear combination of cycles with different norms and phases. Just as with the choice of the clustering method using the C_ω eigenfunctions, the choice of the cycle enumeration algorithm when using S_ω's eigenfunctions could take many possible forms. We call our choice the *co-conformal cycle hunt*.

To describe the co-conformal cycle hunt we first must describe a co-conformal transformation. Recall, conformal deforms the rate one laves the vertices. A co-conformal transformation in essence magnetizes the vertices. Given a vector of positive weights α the co-conformal transformation of P determined by α produces the discrete chain

$$Q = [P[\alpha]\mathbf{1}]^{-1}P[\alpha].$$

Notice

$$\frac{q_{ij}}{q_{ik}} = \frac{\alpha_j}{\alpha_k}\frac{p_{ij}}{p_{ik}}$$

which tells us that the larger the relative size of α at j to k then the stronger the tendency to choose j over k whenever the choice arises.

To state the algorithm, notice each eigenfunction determines a probability measure $||f_\omega||^2$ and we can use this measure to co-conformally deform our chain via $\alpha = g(||f_\omega||^2) > p0$. For each point j iterate finding the point i with the maximal (i, j) value of Q and stopping if this value is less than or equal to $q0$ or you find a loop. Collected all the $(q0, p0)$ cycles of the first K eigen-function.

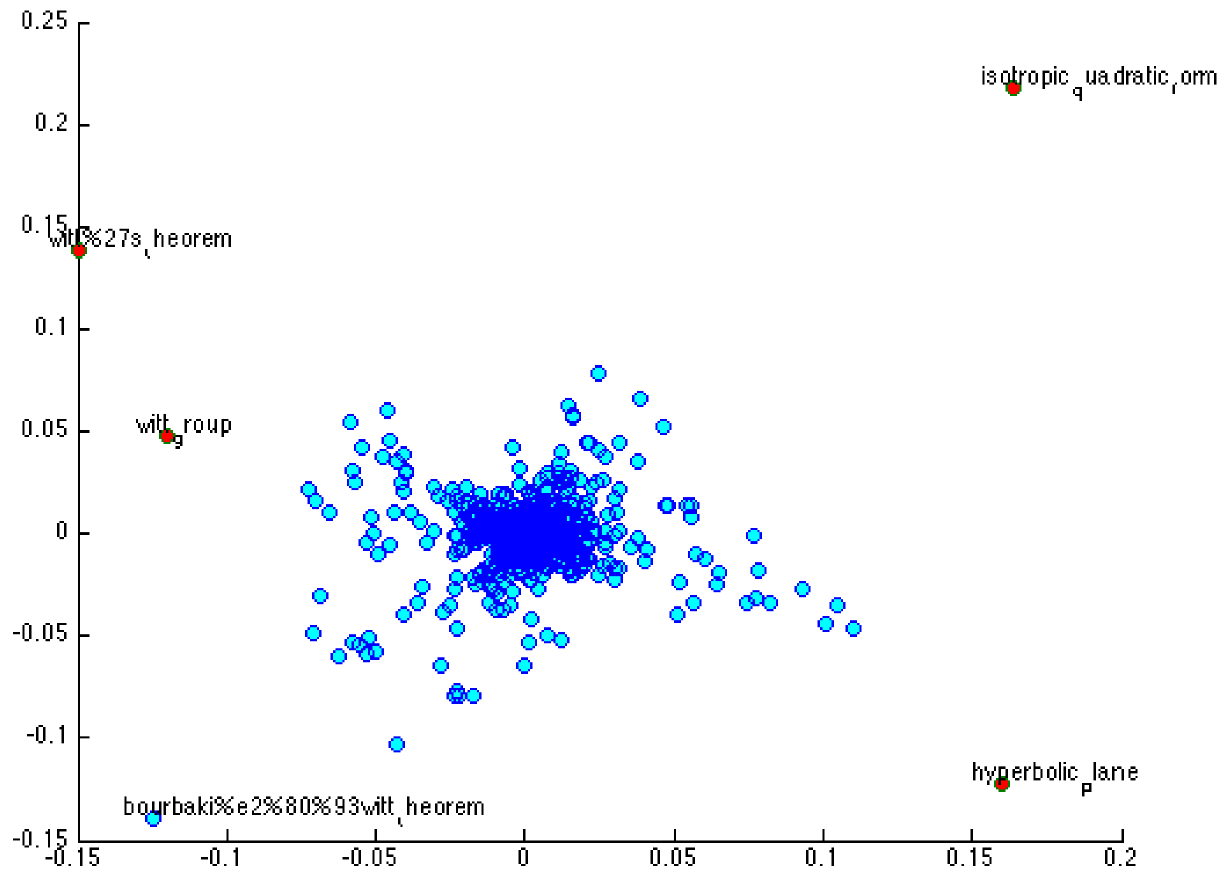

Figure 6: A eigenfunction of the S_ω operator.

Embedded Cycles Example: Here we construct a family of random examples to explore how cycles and there detection behave. We choose to produce a model were we add to $\varepsilon(rand(N,N))$ independent cycles of size $[C_1,...,C_K]$. In the exercises, you are asked to do this and explore the resulting eigenvalue distributions and the The Conformal Cycle Hunt. The example in the figure was from an $N = 1000$ run with $[C_1,...,C_K] = [3,4,5,7,8,11,12]$ and $\varepsilon = 0.1$. In figure 7 we see an eigenfunction with the cycles found by the conformal cycle hunt algorithm corresponding to those divisible by 4 of sizes $[4,5,8,12]$. The co-conformal cycle hunt algorithm has no trouble identifying all the and only the specified cycles.

Exercises:

1. For a fixed cycle chain of length N, determine the κ and the eigenfunctions.

2. Our choice of the cycle scenario here is perhaps the simplest scenario. But any operator will do,

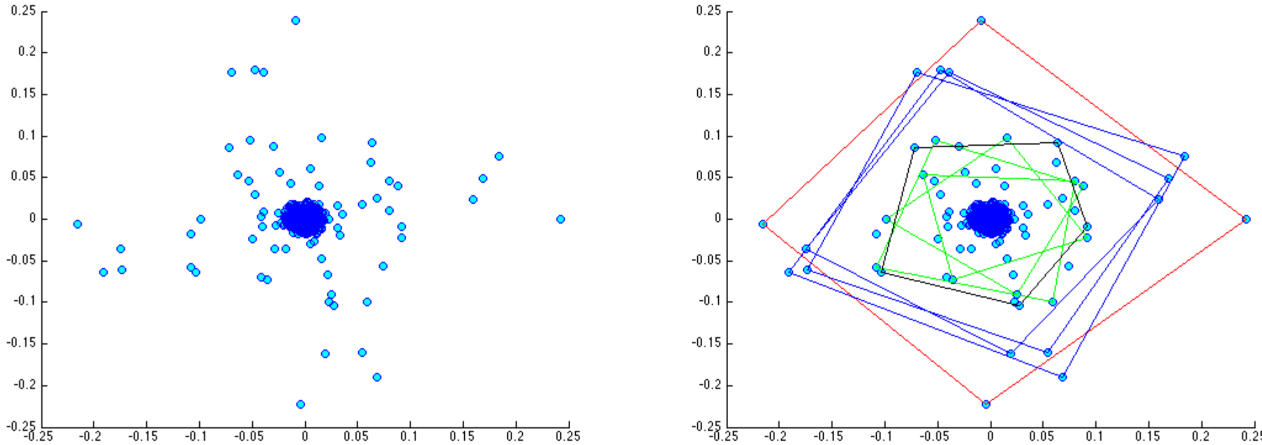

Figure 7: Here witness that it is necessary to have an algorithm to pick the cycles out of the eigenfunctions. We use the co-conformal cycle hunt.

though one that is highly localized would be most natural. Develop a second scenario.

3. Prove or disprove the co-conformal transformation preserve reversibility? (Extra Credit: Write a paragraph describing the answer is remarkable.)

4 Scenarios in context

In section 2 we saw that the context operator C_ω determined the reversible aspects of the geometry while in section 3 we saw that scenario operator S_ω determined the non-reversible aspects of the geometry. What is left is to explore the interaction between the two, which amounts to studying how the context and scenarios couple. This coupling will helps us finds anomalies, namely scenarios that are out of context. And to accomplish this we will take a decidedly quantum mechanical approach viewing our Hermitian operators C_ω and S_ω as measurement operators. We adopt the usual shorthand

$$\langle \psi, A\psi \rangle_\omega = \langle A \rangle_\psi$$

which is the expected value of the A eigevalues using the distribution determined by ψ. Letting $\hat{A}_\omega = A_\omega - \langle A_\omega \rangle_\psi$, the variance of the eigenvlaues is $\langle \hat{A}^2 \rangle_\psi$, and we denote the standard deviation as $\sigma_\psi A = \sqrt{\langle \hat{A}^2 \rangle_\psi}$. If we are given a state that captures some aspect of the geometry, call it ψ, then if the geometry is coupled to both the scenarios and context then $\sigma_\psi C_\omega$ and $\sigma_\psi S_\omega$ will be small, hence $(\sigma_\psi C_\omega)(\sigma_\psi S_\omega)$ could be used to capture the simultaneous ability of C_ω and S_ω to capture ψ. Better yet we have the RobertsonSchrdinger inequality in exercise 3, and the even more basic lemma:

The Uncertainty Principle for Markov Chains Lemma:

$$\left(\sigma_\psi C_\omega\right)\left(\sigma_\psi S_\omega\right) \geq \left|\frac{1}{2}\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_\psi\right|$$

proof: Notice

$$[C_\omega, S_\omega] = C_\omega S_\omega - S_\omega C_\omega = i[\Delta_\omega, \Delta_\omega^*],$$

so

$$\begin{aligned}
\left|\tfrac{1}{2}\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_\psi\right| &= \left|\tfrac{1}{2}\left\langle[C_\omega, C_\omega]\right\rangle_\psi\right| && \text{(def. plus foil)}\\
&= \left|\tfrac{1}{2}\left\langle[\hat{C}_\omega, \hat{S}_\omega]\right\rangle_\psi\right| && \text{(differ by constants)}\\
&= \tfrac{1}{2}\left|\left\langle\psi, \hat{C}_\omega\hat{S}_\omega - \hat{C}_\omega\hat{S}_\omega\psi\right\rangle\right| && \text{(definition)}\\
&= \tfrac{1}{2}\left|\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle - \left\langle\hat{S}_\omega\psi, \hat{C}_\omega\psi\right\rangle\right| && \text{(self adjoint)}\\
&= \left|Im\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle\right| && \text{(Hemitian Inner-product)}\\
&\leq \left|\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle\right| && (|a+ib| = \sqrt{a^2+b^2})\\
&\leq \sqrt{\left\langle\hat{C}_\omega\psi, \hat{C}_\omega\psi\right\rangle}\sqrt{\left\langle\hat{S}_\omega\psi, \hat{S}_\omega\psi\right\rangle} && \text{(Cauchy-Schwarz inequality)}\\
&= \sqrt{\left\langle\psi, \hat{C}_\omega^2\psi\right\rangle}\sqrt{\left\langle\psi, \hat{S}_\omega^2\psi\right\rangle} && \text{(Self adjoint)}\\
&= \left(\sigma_\psi C_\omega\right)\left(\sigma_\psi S_\omega\right) && \text{(defintion)}
\end{aligned}$$

Q.E.D

So the joint geometry is lurking in $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_\psi\right|$, and geometry tightly coupled to both the context and scenarios will have a large $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_\psi\right|$, while geometry not coupled to both with have a small $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_\psi\right|$. Hecne we call $[\Delta_\omega, \Delta_\omega^*]$ the anomaly operator in this case. Let us witness what this all means in an example (and explore further what this means in the exercises).

Quantum Toy: In figure 8, we see our Toy model with the labeled cycles adjoined (all with =weight 11). Note the blue ones are clearly coupled to the geometry while the black ones are not. For the lower blue cycle $C_{l,b}$ we have $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_{\chi_{C_{l,b}}}\right| = 0.0039$, for the upper blue cycle $C_{u,b}$ we have $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_{\chi_{C_{u,b}}}\right| = 0.0141$. While the left black cycle $C_{l,k}$ we have $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_{\chi_{C_{l,k}}}\right| = 0.5901$ and for the right black cycle $C_{r,k}$ we have $\left|\left\langle[\Delta_\omega, \Delta_\omega^*]\right\rangle_{\chi_{C_{r,k}}}\right| = 0.3157$. This order of magnitude increase is the sort of thing we expect from the Uncertainty Principle for Markov Chains.

Exercises:

1. Construct a random chain with *Clus* number of clusters and *Cyc* cycles. Chose the cycle to be random placed **inside** clusters, and also chosen randomly independent of the clusters. Show the eigenvalues distribution of C_ω and S_ω and nearly identical in each case, but the the distribution of the eigenvalues of $[\Delta_\omega, \Delta_\omega^*]$ are entirely different.

2. Prove

$$\left.\frac{d}{dt^2}e^{t\Delta^*}e^{t\Delta}e^{t\Delta^*}e^{t\Delta}\right|_{t=0} = [\Delta, \Delta^*]$$

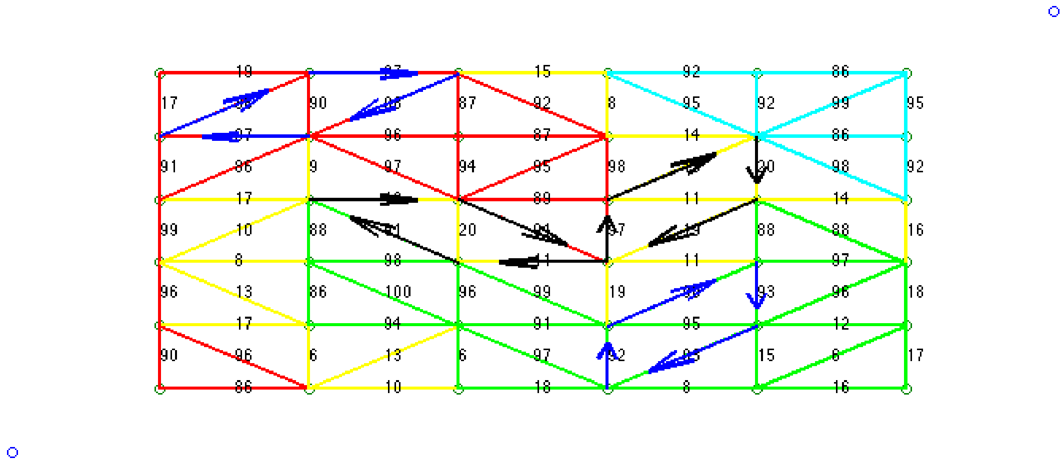

Figure 8: Anomalies waiting to be detected.

.

3. We can improve our estimate by literally computing the σ literally or using the full Robertson-Schrdinger. Prove

$$(\sigma_\psi C_\omega)\left(\sigma_\psi S_\omega\right) \geq \sqrt{\frac{1}{4}\left|\langle[\Delta_\omega,\Delta^*_\omega]\rangle_\psi\right| + \frac{1}{4}\left|\langle\{\hat{\Delta}_\omega,\hat{\Delta}^*_\omega\}\rangle_\psi\right|}$$

where $\{A,B\} = AB + BA$.

4. Interpret the solutions to $\frac{d\psi}{dt} = ihC_\omega\psi$ and $\frac{d\psi}{dt} = ihS_\omega\psi$.

5. Now if A is a cluster $\left|\langle[\Delta_\omega,\Delta^*_\omega]\rangle_{\chi_A}\right|$ will be small if the space's cycles are coupled to this cluster and large if not. Construct's some examples and verify this principle.

References

[1] Ashok K. Chandra , Prabhakar Raghavan , Walter L. Ruzzo , Roman Smolensky, and Prasoon Tiwari, *The Electrical Resistance Of A Graph Captures Its Commute And Cover Times*, http://www.cs.washington.edu/homes/ruzzo/papers/resist.pdfDBLP

[2] Peter G. Doyle and J. Laurie Snell, *Random Walks and Electric Networks*, http://arxiv.org/abs/math/0001057

[3] Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability*, `http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf`

[4] Peter G. Doyle, Gregory Leibon, and Jean Steiner *Conformal geometry of Markov chains*, pickle.pdf as posted on CSSS10 Wiki

[5] Peter G. Doyle, Gregory Leibon, and Jean Steiner *Markov chains in a shoebox* , shoe.pdf as posted on CSSS10 Wiki

[6] Gregory Leibon, Scott Pauls, Dan Rockmore, and Robert Savell *Statistical Learning for Complex Systems an Example-driven Introduction* , book.pdf as posted on CSSS10 Wiki

[7] Kyle Lad *Python code to mine the Space of Mathematics* , SpaceOfMath.zip as posted on CSSS10 Wiki