mean-field probabilistic analysis
of combinatorial optimization problems

max-type
RDEs

local weak convergence
of random networks

Each is a small corner of some large subject.

Input to TSP (traveling salesman problem) is usually taken as n points in R^2. What actually matters is the $\binom{n}{2}$ interpoint distances (costs); not necessarily proportional to Euclidean distance.

For probabilistic analysis of TSP (and related problems: MST = minimum spanning tree; MM = minimum matching) we want some model of $\binom{n}{2}$ random interpoint distances.

To compare different models, need a **scaling convention**: distance from point to nearest neighbor is order 1 as $n \to \infty$.

Uniform-Euclidean model: n uniform random points in square of area n; Euclidean distance.

simulation picture

$L_n :=$ length of TSP tour on n random points in $[0, \sqrt{n}] \times [0, \sqrt{n}]$.

What is known about L_n?

$$n^{-1} L_n \to c$$

(Beardwood - Halton - Hammersley 1959, Kingman 1973).

$$\text{var } (L_n) \leq c' n$$

(Steele 1981; corresponding large deviation bounds Rhee - Talagrand 1989)

What is <u>not</u> known?

- value of c
- CLT for L_n?
- sample path properties of optimal tour?

Two other models for interpoint distances over n points.

Disordered lattice. Take a box $[0, n^{1/d}]^d$ in the lattice Z^d; to each edge e assign independent random lengths ξ_e.

Mean field model, pseudo-dimension d. Take all $\binom{n}{2}$ distances to be independent, distributed as $n^{1/d}\xi$,

$$P(\xi > x) = \exp(-x^d).$$

This mimics the distances in uniform-Euclidean model from a given point to other points.

<u>Observation</u>: in each of these 3 models we can let $n \to \infty$ and get an infinite-point model.

For the disordered lattice, limit is just disordered Z^d.

For uniform-Euclidean, the limit is just Poisson-Euclidean: a Poisson process of points on R^d with Euclidean distances.

For the mean-field model, the limit is PWIT$_{(d)}$ **(Poisson weighted infinite tree)**.

Root has infinite number of children; edge lengths are realization of a Poisson process $(\xi_i^r,\ 1 \le i < \infty)$ on $(0, \infty)$ with mean function x^d. Recursively, each vertex v has infinite number of children; edge lengths are realization of a Poisson process $(\xi_i^v,\ 1 \le i < \infty)$ on $(0, \infty)$ with mean function x^d.

.

PROGRAM

Study solutions of size-n CO problems under these 3 (and other?) probability models by arguing that the solutions converge as $n \to \infty$ to the solution of the corresponding CO problem in the infinite-point model.

(We have parted company with classical CO; worst-case analysis of infinite problems doesn't make much sense; we're not directly interested in algorithmic questions).

3 issues:

Conceptual: what does "solution of an infinite CO problem" mean? What does convergence mean?

Calculational: can we find optimal solution to infinite random CO problem?

Technical: proving convergence.

Local weak convergence and stationary random networks (Aldous - Steele survey 2002)

Network: graph with edge lengths; assume locally finite. For finite or infinite rooted networks, there is a natural notion of convergence: local convergence.

Given n-vertex random networks G_n, take uniform random root X_n and ask if the rooted networks $G_n[X_n]$ converge in distribution to some limit rooted infinite network G_∞. This is local weak convergence of G_n to G_∞. Has arisen in several contexts.

- uniform random n-trees (Grimmett 1980)
- general models of random n-trees (Aldous 1991)
- random planar lattices/triangulations (Chassaing-Schaffer 2002; Angel-Schramm 2002)
- mean-field model and PWIT (Aldous 1992)
- graph spectra (Grigorchuk-Zuk 1999)
- RW recurrence on planar graphs (Benjamini-Schramm 2001)

A rooted infinite network G_∞ arising as a local weak limit automatically inherits a certain "stationarity" property: <u>involution invariance</u> (II). Slightly fussy to define in general; when G_∞ is regular, II is the property that re-rooting at unform random neighbor of the root preserves the distribution.

Central math idea: given an infinite II rooted random network G_∞, and given a suitable CO problem, one can define a constant $c(\text{problem}, G_\infty)$, illustrated for TSP below.

A feasible solution on a realization of G_∞ is an edge-set E_∞ which is a union of vertex-disjoint doubly-infinite paths covering every vertex. Require (G_∞, E_∞) to be jointly II. Define

$$c(TSP, G_\infty) = \inf \tfrac{1}{2} E(\text{length } (e_1) + \text{length } (e_2))$$

where e_1, e_2 are the edges of E_∞ at the root, and the *inf* is over all jointly II choices of (G_∞, E_∞).

To motivate the definition, we digress to look at minimum spanning forests (MSFs). Consider an infinite network with
(*) distinct edge-lengths.

Explicit construction of the MSF:
$e = (v_1, v_2) \in MSF$ iff, in the subgraph in the subgraph with only edges of length $<$ length (e), the vertices v_1 and v_2 are in distinct components, at least one of which is finite.

Abstract set-up. On an infinite network, say a feasible solution is an edge-set E_∞ which is a union of vertex-disjoint infinite trees, covering every vertex. Given II random G_∞, require (G_∞, E_∞) to be jointly II. Define
$c(MSF, G_\infty) = \inf$

$\frac{1}{2} E \sum \{\text{length } (e) : \ e \in E_\infty, \ e \text{ incident to root}\}$

where the *inf* is over all jointly II choices of (G_∞, E_∞).

Theorem. For II random G_∞ satisfying (*), the *inf* is attained uniquely by the MSF given by the explicit construction.

MST Convergence theorem. Consider n-vertex connected networks G_n, such that $G_n[X_n]$ converges in distribution to some G_∞ satisfying (*). Then jointly with the above convergence we have

$$MST(G_n) \ \stackrel{d}{\to} \ MSF(G_\infty).$$

And (under uniform integrability of edge-lengths of MST(G_n))

$$n^{-1} \ E \text{ length } MST(G_n) \to c(MSF, G_\infty).$$

These are "soft" results, except (perhaps?) for one point. Easy to show that, in an II infinite random network,

$$E(\text{degree of root in MSF}) \geq 2.$$

We don't know a direct proof of "$= 2$"; our indirect proof uses a semi-hard

Theorem. Given any infinite II random network G_∞, there exist n-vertex networks G_n such that $G_n[X_n] \xrightarrow{d} G_\infty$.

To see why not obvious, consider the special case where G_∞ is the non-random infinite rooted binary tree.

This is all some new corner of classical measure-theoretic ergodic theory.

The MST Convergence Theorem applies to all 3 models. For disorder Z^d and Poisson-Euclidean, the limit $c(MSF, \cdot)$ involves non-explicitly-known percolation functions.

Reinterpret Frieze (1985) as saying

$$c(MSF, \mathrm{PWIT}_{(d=1)}) = \zeta(3) := \sum_{i \geq 1} i^{-3}$$

and then Timofeev (1988) as

$$c(MSF, \mathrm{PWIT}_{(d)}) = d^{-1} \sum_{i \geq 0} \frac{(i-1+1/d)!}{i!(i+1)^{2+1/d}}.$$

It is unknown whether there is a general TSP convergence theorem.

(?) if we have local weak convergence of G_n to G_∞, if edge-lengths of $TSP(G_n)$ are UI, then

$$\lim_n n^{-1} E \text{ length } \ TSP(G_n) = c(TSP, G_\infty).$$

For disordered Z^d and Poisson-Euclidean, this is true: "boundary effect negligible". But can't hope to calculate limit explicitly in those models.

The mean-field model is interesting, because of the replica method analysis by Mézard - Parisi (1986). From our viewpoint, what they did is to argue

$$c(TSP, \text{PWIT}_{(d=1)}) = 2.04....$$

This limit constant arises as follows.

$$c = \int_0^\infty \ xP(X_1 + X_2 > x) \ dx$$

where X_1, X_2 are independent with the distribution determined as the solution of the RDE (recursive distributional equation)

$$X \ \stackrel{d}{=} \ \min{}_i^{[2]}(\xi_i - X_i)$$

where $0 < \xi_1 < \xi_2 < \ldots$ are the points of a Poisson(rate 1) process, and $\min^{[2]}$ denotes second-minimum.

Distinguished edge e^* of PWIT specifies two rooted subtrees $\mathbf{T}^{right}, \mathbf{T}^{left}$.

Optimal solution either uses/doesn't use e^*; restriction of solution to $\mathbf{T}^{right}$ the root has either degree 1 or degree 2. So on $\mathbf{T}^{right}$ define heuristically

$$\begin{aligned} X^{right} &= \text{total length opt. solution, rootdegree } = 2 \\ &- \text{total length opt. solution, rootdegree } = 1. \end{aligned}$$

Optimal solution on $\mathbf{T}$ will use e^* iff

$$\text{length } (e^*) \leq X^{right} + X^{left}. \quad \text{[inclusion criterion]}.$$

At a given vertex, the mean density of edges with lengths $\in [x, x + dx]$ equals dx; so contribution to mean length of TSP equals

$$\int_0^\infty x \ P(x \leq X^{right} + X^{left}) \ dx.$$

Natural recursive structure of $\mathbf{T}^{right}$ leads to the RDE for distribution of X.

In the similar minimum matching (random assignment) problem, Aldous (2001) goes through the details of a rigorous proof. Adapting to mean-field TSP involves two issues

- existence/uniqueness of solution of the RDE

$$X \stackrel{d}{=} \min_i^{[2]}(\xi_i - X_i)$$

- constructing feasible TSP solutions in finite-n mean-field model which mimic optimal solution on PWIT.

In MM, second part was Horrible Hack. More elegant route would be via

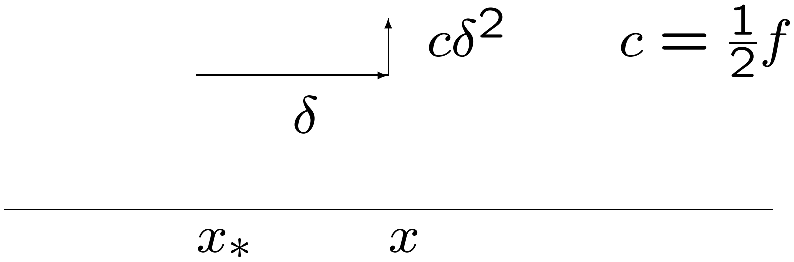
Synthesis Conjecture. Given local weak convergence of G_n to G_∞, and given extra structure (e.g. edge-set E_∞) such that (G_∞, E_∞) is jointly II, we can construct analogous extra structure E_n on G_n such that we have joint local weak convergence of (G_n, E_n) to (G_∞, E_∞).

Alas, I guess false!

lotsa other RDEs - survey table

Calculus meets Combinatorial Optimization

A minimum x_* satisfies $f'(x_*) = 0, \; f''(x_*) > 0$.



$c = \frac{1}{2} f''(x_*)$

In statistical physics language: typically the **scaling exponent** $= 2$ in all dimensions. That is, for $f : R^d \to R$ and a minimum x_*,

$$\min_{x:|x-x_*|=\delta} (f(x) - f(x_*)) \asymp \delta^2.$$

Consider a TSP or MST or . . . problem. Compare an arbitrary feasible solution $\mathbf{x}$ with the optimal (minimal) solution $\mathbf{x}_*$ by the two quantities

$\delta_n(\mathbf{x}) = \frac{1}{n}\times$ (number of edges in $\mathbf{x}$ but not in $\mathbf{x}_*$)

$\varepsilon_n(\mathbf{x}) = \frac{1}{s(n)}\times$ (cost difference between $\mathbf{x}$ and $\mathbf{x}_*$)

where $s(n)$ is the growth rate for optimal cost. Now define $\varepsilon_n(\delta)$ to be the minimum value of $\varepsilon_n(\mathbf{x})$ over all feasible solutions $\mathbf{x}$ for which $\delta_n(\mathbf{x}) \geq \delta$. Although the function $\varepsilon_n(\delta)$ will depend on n and the problem instance, we anticipate that for typical instances drawn from a suitable probability model it will converge in the $n \to \infty$ limit to some deterministic function $\varepsilon(\delta)$. The *universality* paradigm suggests there might be scaling exponents α defined by

$$\varepsilon(\delta) \asymp \delta^\alpha \text{ as } \delta \downarrow 0$$

and that the exponents should be robust under model details and should serve to categorize different combinatorial optimization problems into a small set of universality classes. If our analogy with freshman calculus is correct, the simplest combinatorial optimization problems should have scaling exponent 2.

This program is work-in-progress. What we believe is

• for MST, for any reasonable model (including the 3 models we have discussed) the scaling exponent $= 2$.

• for TSP in the mean-field model of distance, the scaling exponent $= 3$.

We will say a little more in a moment.

Understanding the scaling exponent for TSP on random points in R^d seems far beyond current theory; but is amenable to simulation.

Percus diagram

MSF has scaling exponent = 2 for a broad class of probability models.

An edge e is in the MSF iff

$$\text{length } (e) \leq \ell^*(e)$$

where $\ell^*(e)$ depends on all the other edge-lengths in a certain way. So an edge e not in the MSF has an <u>excess length</u>

$$\text{excess}(e) := \text{length } (e) - \ell^*(e) > 0.$$

Suppose, in a random II infinite network, that for a random edge $e \notin MSF$ there is a density function $f(\cdot)$ on $(0, \infty)$ for the excess. The condition $0 < f(0+) < \infty$ is essentially sufficient for the scaling exponent to equal 2.

Why? Can construct a near-optimal SF by taking an edge with excess $\le \beta$, adding this edge, deleting the longest other edge in the cycle. Repeat for all edges with excess $\le \beta$. Then

$$\begin{aligned}
\delta(\beta) &:= \text{proportion of edges of MSF that} \\
& \text{are not in near-opt SF} \\
&\asymp \beta \\
\varepsilon(\beta) &:= \text{extra length per vertex} \\
& \text{of near-opt SF over MSF} \\
&\asymp \beta^2.
\end{aligned}$$

No other construction does better.

The condition $f(0+) \in (0, \infty)$ easily checked for PWIT and for disordered Z^d; intuitively clear in Poisson-Euclidean model but some technical issues of proof ....

Mean-field probability model of TSP has scaling exponent = 3.

One can give a (heuristic) <u>exact</u> calculation! Issue is a "penalized TSP" where a specified proportion of edges must be distinct from optimal TSP edges. Introduce a Langrange multiplier λ; effect is to penalize edges in optimal tour by increasing their lengths by λ. Seek to study relation between

$$\begin{aligned} \delta(\lambda) &:= \text{proportion of edges in penalized-opt} \\ &\quad\ \text{tour which are not in opt tour} \\ \varepsilon(\lambda) &:= \text{extra length per vertex of penalized-opt} \\ &\quad\ \text{tour over opt tour.} \end{aligned}$$

.

Recall X defined as

$$\begin{array}{rl} & \text{total length opt. solution, rootdegree } = 2 \\ - & \text{total length opt solution, rootdegree } = 1 \end{array}$$

using original edge lengths. Define Y and Z in the same way using <u>penalized</u> edge lengths; Y (resp. Z) is case where in optimal TSP we require root degree = 2 (resp. = 1).

Get an inclusion criterion: distinguished edge e^* included in penalized-opt solution if

$$\begin{array}{rcl} \text{length } (e^*) + \lambda & < & Z^{left} + Z^{right} \ (e^* \in \text{opt TSP}) \\ \text{length } (e^*) & < & Y^{left} + Y^{right} \ (e^* \notin \text{opt TSP}). \end{array}$$

Recursive structure of $\mathbf{T}^{right}$ leads to a RDE for the triple (X, Y, Z).

$$\begin{aligned}
X &= \min_i{}^{[2]}(\xi_i - X_i) \\
I &= \arg\min_i{}^{[2]}(\xi_i - X_i) \\
Z &= \min_i{}^{[2]}(\xi_i - Y_i) \\
Y &= \min_i{}^{[2]}\left(\xi_i - Y_i 1_{(i \neq I)} - (Z_i + \lambda) 1_{(i=I)}\right).
\end{aligned}$$

We end up with explicit formulas for $\delta(\lambda)$ and $\varepsilon(\lambda)$ in terms of two copies of solutions (X, Y, Z) of the RDE above.

$$\begin{aligned}
\delta(\lambda) &= \int_0^\infty P(X^{(1)} + X^{(2)} < x < Y^{(1)} + Y^{(2)})\ dx \\
\varepsilon(\lambda) &= \text{[more complicated formula]}
\end{aligned}$$

Can solve by numerical or Monte Carlo methods. Solution shows good fit to $\varepsilon \asymp \delta^3$.

table.