Euclidean Stacker-Crane: Approximate Solution

Daniel Friedman, Sharath Raghvendra

CS Department, Virginia Tech

October 8, 2015

Common Problem

Suppose you're the last taxi driver on Earth. You're very busy, so people tell you their pickup and destination at the beginning of the day. In what order should you serve individuals so as to be most fuel efficient?

Formal Definition

Definition (Stochastic Euclidean Stacker-Crane Problem)

Let $(x_1, y_1), \ldots, (x_n, y_n)$ be pickup-delivery pairs in $[0, 1]^2$ sampled from a distribution with PDF f. Find the minimum-length tour through all points, preserving the pickup-delivery assignment.

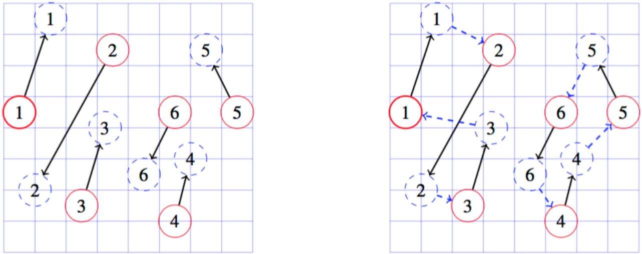
In other words, find assignments (y_i, x_j) so that the distance over all is minimized.

Previous Work

- Solving for exact solutions is NP-Hard
- [Treleaven et al] found an $O(n^{2+\epsilon})$ asymptotically optimal solution for uniformly distributed points, i.e.,

$$\lim_{n\to\infty} \frac{\mathbb{E}[\text{Cost of algorithmic solution on } n \text{ pairs}]}{\mathbb{E}[\text{Cost of exact solution on } n \text{ pairs}]} = 1$$

SPLICE (Matching)



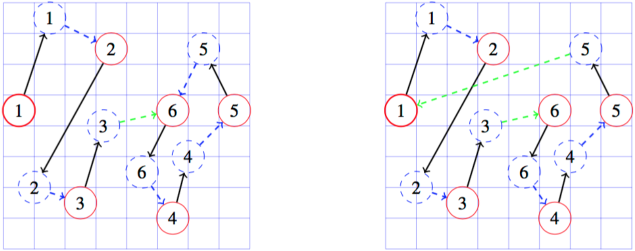
(a) Line 2: 6 pickup-to-delivery links are added.

(b) Line 3: 6 matching links are added. The number of disconnected subtours is $N = 2$.

Figure: Euclidean Bipartitide Matching

[Treleaven et al]

SPLICE (Merging)



(c) Line 10. Algorithm state: `prev` = `base` = 3, $k = 1$. The link $y_3 \rightarrow x_1$ is removed, `next` is assigned the value 6, the link $y_3 \rightarrow x_6$ is added, `prev` is assigned the value 5.

(d) Line 15. Algorithm state: `prev` = 5, `base` = 3. The link $y_5 \rightarrow x_1$ is added an the tour is completed.

Figure: Merging tours

[Treleaven et al]

Observations

- Most EBMP solvers use the Hungarian Algorithm which runs in $O(n^3)$, the $O(n^{2+\epsilon})$ is theoretical
- Only solves for uniformly distributed samples
- Takes about 20 minutes to solve for 400 pairs

Near-SPLICE Algorithm

- Assumes continuous PDF on $[0,1]^2$
- Linear-time algorithm
- Asymptotically near-optimal, i.e., given $\epsilon > 0$ we can choose parameters so that

$$\lim_{n\to\infty} \frac{\mathbb{E}[\text{Cost of algorithmic solution on } n \text{ pairs}]}{\mathbb{E}[\text{Cost of exact solution on } n \text{ pairs}]} = 1 + \epsilon$$

- Easier implementation

Spatial Partitioning Tree

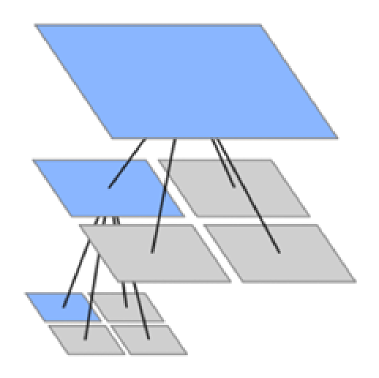


Figure: Quad Tree Partitioning

http://www.thelowlyprogrammer.com/2011/05

PQT

Definition

A Probabilistic Quad-Tree (PQT) is a spatial partitioning tree. It is formed so that every node either has 4 children that subdivide the area corresponding to the node, or it is a leaf whose probability over the corresponding area is less than some fixed p. Each leaf is therefore a cell $\mathcal{C}_i$ with side-length $\ell_{\mathcal{C}_i}$ and probability $p_{\mathcal{C}_i} = Pr(\mathcal{C}_i)$ for $i = 1, \ldots, N$.

Example

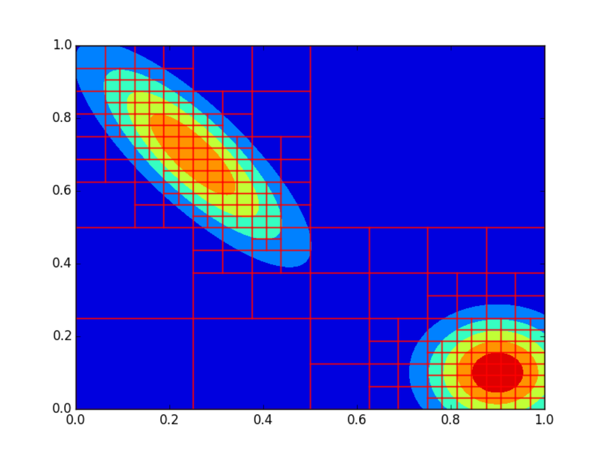


Figure: PQT over Restricted GMM with $p = 0.01$

Expected Length

Definition

Let f be a PDF on a space V and norm $\|\cdot\|$. The V-line picking is the expected distance between two points sampled from V according to f. Assuming points are sampled i.i.d. this is given by

$$\mathbb{E}[\ell_f] = \iint_{V^2} \|p_1 - p_2\| f(p_1) f(p_2) dp_1 dp_2$$

Special Case

In the special case of a uniform distribution, $f(x, y) = 1$,

$$\mathbb{E}[\ell_f] = \frac{1}{15}\left[\sqrt{2} + 2 + 5\ln(1+\sqrt{2})\right] \approx 0.5214$$

Expected Excess

Definition

The excess random variable is given by $\mathcal{E}_n = |\sum_i X_i - \sum_i Y_i|$ where $\{X_i\}$ and $\{Y_i\}$ are sets of Bernoulli random variables with probability p.

Theorem

For expected excess:

$$\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{E}_n]}{\sqrt{n}} = 2\sqrt{\frac{p(1-p)}{\pi}}$$

Random Permutation

Theorem

If $\sigma \in S_n$ *and* $Pr[\sigma] = 1/n!$, *then*

$$\mathbb{E}[\text{\# cycles in } \sigma] = H_n \approx \log(n)$$

Near-SPLICE

1. Form PQT over the unit square
2. Connect each available delivery to an available pickup within each cell
3. Connect each available delivery to an available pickup between cells
4. Merge tours

Example

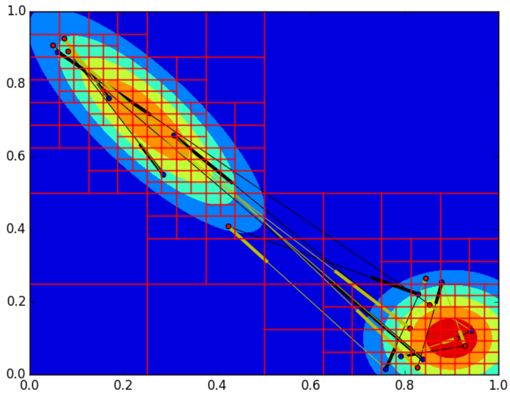


Figure: Near-SPLICE with 10 pairs

Example

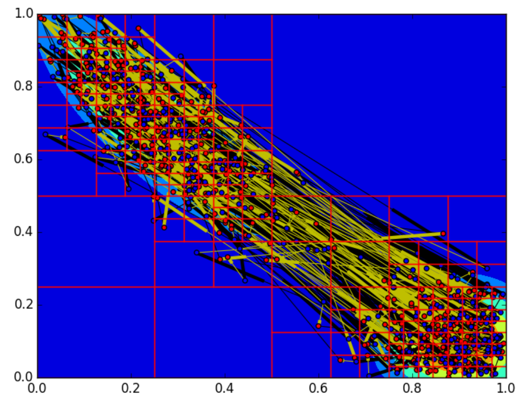


Figure: Near-SPLICE with 400 pairs

Analysis

Denote:

- Expected cost of a pairs within cells: $\mathbb{E}[\ell_{\mathcal{C}}]$
- Expected length on entire grid: $\mathbb{E}[\ell_f]$
- Expected number of tours: $\mathbb{E}[\tau_n]$
- Expected Excess: $\mathbb{E}[\mathcal{E}_n]$
- Maximum value of the PDF: f_M
- Expected cost ratio: $\mathbb{E}[C_n]$

Number of cell bound

f is continuous and therefore obtains a maximum f_M on the unit square. By creating a grid with side-length p/f_M then all cells have probability $\leq p$, so

$$N \leq \frac{f_M}{p}$$

NB. this also means that a PQT may be replaced by a fine grid

Expected cost of pairs within cells

The maximum distance in a cell is $\sqrt{2}\ell_{\mathcal{C}_i}$, so

$$\mathbb{E}[\ell_{\mathcal{C}}] \leq \sum_{i=1}^{N} \sqrt{2}\ell_{\mathcal{C}_i} Pr(\mathcal{C}_i) \leq \sqrt{2\left(\sum_{i=1}^{N} \ell_{\mathcal{C}_i}^2\right)\left(\sum_{i=1}^{N} p_i^2\right)} \leq \sqrt{2Np} \leq \sqrt{2f_M p}$$

Expected Competitive Ratio

$$\mathbb{E}[C_n] \leq \frac{n\,\mathbb{E}[\ell_f] + n\sqrt{2 f_M p} + \frac{1}{\sqrt{2}}\,\mathbb{E}[\mathcal{E}_n] + \sqrt{2}\,\mathbb{E}[\tau_n]}{n\,\mathbb{E}[\ell_f]}$$

Asymptotic Expected Competitive Ratio

$$\lim_{n \to \infty} \mathbb{E}[C_n] \leq 1 + \frac{\sqrt{2 f_M p}}{\mathbb{E}[\ell_f]}$$

For any $\epsilon > 0$, if we choose $p < \frac{\epsilon^2 \, \mathbb{E}[\ell_f]^2}{2 f_M}$, then the asymptotic competitive ratio is $1 + \epsilon$.

Further Work

- Different distributions for pickup and deliveries
- Better approximation for the maximum number of cells in a PQT
- Extend to multiple taxis

References

- “An Asymptotically Optimal Algorithm for Pickup and Delivery Problems” - Treleaven et al
- http://www.thelowlyprogrammer.com/2011/05/game-of-life-part-2-hashlife.html