10/5 Gossip Algorithms

Thursday, October 05, 2006
9:56 AM

Graph $G = (V, E)$

vector of values $x_t(i)$ = value at i at time t

At each time, two nodes communicate

If (i, j) communicate at time t

$$x_{t+1}(k) = \begin{cases} \dfrac{x_t(i) + x_t(j)}{2} & \text{if } k \in \{i, j\} \\ x_t(k) & \text{o.w.} \end{cases}$$

Let p_{ij} be the prob that (i, j) communicate in a round, and assume p_{ij} fixed for all time.

Can view p_{ij} as supplying a wted graph.

Assume it is connected.

Claim: this procedure probably converges.

Converges to average, as

$$\sum_i x_t(i) = \sum_i x_{t+1}(i),$$

so sum stays fixed.

let $X_{ave} = \frac{1}{n} \sum_i x_0(i)$

Observe $X_{ave} \mathbb{1}$ is fixed point

Want to show is only fixed point

Want to bound how long it takes to converge.

State using linear algebra.

let W_{ij} be the matrix

$$W_{ij} = I - \frac{(e_i - e_j)(e_i - e_j)^T}{2}$$

where $e_i = (0, 0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0)$

Note $(e_1 - e_2)(e_1 - e_2)^T = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

so, $(e_i - e_j)(e_i - e_j)^T$ is Laplacian matrix of graph with just one edge from i to j.

$I - \frac{1}{2}(e_1 - e_2)(e_1 - e_2)^T$

$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & \frac{1}{2} & & 0 \\ & & 1 & \\ 0 & & & 1 \end{pmatrix} \quad \text{etc}$$

If i and j comm at time t,
then

$$x_{t+1} = W_{ij} x_t$$

Note $W_{ij} \mathbb{1} = \mathbb{1}$

Let W_t be matrix used at time t.

Then $x_t = W_t W_{t-1} \cdots W_1 x_0$

<u>Lem</u> if graph of P is connected, then $E[x_t] = x_{ave} \mathbb{1}$

---

1. Recall for a vector v, $E[v] = (E[v(1)], E[v(2)], \ldots E[v(n)])$

2. If x and y are indep vars, $E[xy] = E[x] E[y]$

3. and $\forall x$ and y, $E[x+y] = E[x] + E[y]$

So, if M is a random matrix, v a random vector, independent, then

$$E[Mv] = E[M] E[v]$$

<u>pf</u> suffices to consider $E[(Mv)(i)]$ by (1)

$$= E\left[\sum_i m_{1,i} v_i\right] = \sum_i E[m_{1,i} v_i] \quad \text{by (3)}$$

$$= \sum_i E[M_{1,i}] E[v_i] \quad \text{by (2)}$$

$$= (E[M] E[v])(1)$$

So, $E[x_t] = E[w_t w_{t-1} \cdots w_1 x_0]$

By prev lemma, $E[x_1] = E[w_1] x_0$

To continue this way, work from other end.

as w_t is indep of $w_{t-1}, \ldots, w_1, x_0$, have

$$= E[w_t] \cdot E[w_{t-1} \cdots w_1 x_0]$$

$$= E[w_t] E[w_{t-1}] \cdots E[w_1] x_0$$

We assumed that each w_k was ident dist, and that

$$E[w_k] = \sum_{i,j} p_{ij} W_{ij}$$

$$= \sum_{i,j} p_{ij} \left(I - \tfrac{1}{2} L_{(i,j)}\right) = I - \tfrac{1}{2} \sum_{i,j} p_{ij} L_{(i,j)}$$

$$= I - \tfrac{1}{2} L_p$$

where L_p is the Laplacian of the graph with wt p_{ij} on edge ij.

As the sum of all entries in p is 1, is easy to show all eigs of L_P lie between 0 and 2

pf. if $Lx = \lambda x$, let $x(i)$ have max value

then $Lx(i) = \sum_{j\sim i} p_{ij}\, x(i) - \sum_{j\sim i} p_{ij} (x(j))$

$$\leq x(i) + x(i) \leq 2x(i)$$

So, all eigs of $I - \frac{1}{2}L_P$ are between 0 and 1

Moreover, if graph of P connected, L has eigenvalue 0 with multiplicity 1 (essentially on hw)

So, letting $\mu_1 \geq \cdots \geq \mu_n$ be eigs of $I - \frac{1}{2}L_P$

$\lambda_1 \leq \quad \leq \lambda_n$ " " L_P

have $\mu_i = 1 - \frac{1}{2}\lambda_i$ $\quad \lambda_1 = 0 \quad \lambda_2 > 0$

$v_1 = \mathbb{1}/\sqrt{n}$

so $\mu_2 < 1 \quad \mu_n \geq 0$

$\Rightarrow$ if $E[W_k] = \bar{W}$

$$(\bar{W})^t x_0 = \sum_i (\mu_i)^t v_i (v_i^T x_0) \to \frac{1}{n}\mathbb{1}\,\mathbb{1}^T x_0 = X_{ave}\mathbb{1}$$

So, converges in expectation

But, does it converge in reality?

ex. if $x = 1$ with prob $\frac{1}{2}$

$= -1$ with prob $\frac{1}{2}$ $\quad E[x] = 0.$

Observe: might as well assume $X_{ave} = 0$

as, adding a const makes no change.

Formally, can play with $y_t = x_t - x_{ave}\mathbb{1}$

and is easy to verify $y_t = W_t y_{t-1}$

To show that $y_t \to 0$, will consider potential function $\|y_t\|^2$.

Will prove something like $E[\|y_t\|^2] \le \beta$, β small

As $\|y_t\|^2 \ge 0$, can apply Markov's inequality which says
"For a non-neg r.v. X,
$Pr[X \ge \varepsilon] \le E[X]/\varepsilon$" $\le \beta/\varepsilon$

if β/ε small, will know unlikely $\|y_t\|^2$ is not smaller than ε

$\|y_t\|^2 = \sum_i y_t^2(i)$

how does this change if i and j communicate.
well

$$\|y_t\|^2 - \|y_{t+1}\|^2 = y_t^2(i) + y_t^2(j) - 2\left(\tfrac{1}{2}(y_t(i) + y_t(j))\right)^2$$

$$= \tfrac{1}{2}\left(y_t(i) - y_t(j)\right)^2 \ge 0$$

So, potential goes down => fixed point unique

$$E\left[\|y_t\|^2 - \|y_{t+1}\|^2 \mid y_t\right] = \tfrac{1}{2}\sum_{ij} p_{ij}\left(y_t(i) - y_t(j)\right)^2$$

$$= \tfrac{1}{2} y_t^T L_p y_t$$

as $\sum_i y_t(i) = 0$, $\quad \geq \frac{1}{2} \lambda_2(L_P)\, y_t^T y_t$

$$= \frac{1}{2} \lambda_2(L_P)\, \|y_t\|^2$$

$$E\left[\|y_{t+1}\|^2 \mid y_t \right] \leq \left(1 - \frac{1}{2}\lambda_2(L_P)\right) \|y_t\|^2$$

$$\Rightarrow E\left[\|y_{t+1}\|^2\right] \leq \left(1 - \frac{1}{2}\lambda_2(L_P)\right) E\left[\|y_t\|^2\right]$$

$$\Rightarrow E\left[\|y_t\|^2\right] \leq \left(1 - \frac{1}{2}\lambda_2(L_P)\right)^t \|y_0\|^2$$

note $(1-\alpha)^{\frac{1}{\alpha}} \leq \frac{1}{e}$

$$\leq e^{-\frac{1}{2}\lambda_2(L_P)\cdot t} \cdot \|y_0\|^2$$

in particular, if $t \geq \frac{2}{\lambda_2(L_P)} \cdot \lg(1/\varepsilon^2)$

Then $E\left[\|y_t\|^2\right] \leq \varepsilon^2 \|y_0\|^2$

and, by Markov's Ineq

With prob at least $1-\varepsilon$, $\|y_t\|^2 \leq \varepsilon \|y_0\|^2$

Example: P is uniform random edge from complete graph.

$L_P = \frac{1}{\binom{n}{2}} L_{K_n}$, $\quad \lambda_2 = \frac{1}{\binom{n}{2}} \cdot n \approx \frac{2}{n}$

So need $n \lg(1/\varepsilon^2)$ iters

If choose from regular $k \times k$ grid, get

$$\lambda_2(P) \simeq \frac{1}{4k^2} \cdot \frac{\text{const}}{k^2} \approx \frac{\text{const}}{k^4} = \frac{\text{const}}{n^2}$$

dilate time by n, and see each node communicates on average n times!

Is too many — should be like diameter.