Social Networks and Social Choice	Lecture Date: September 21 and 23, 2010

Lecture: Some non-martingale learning models

Lecturer: Elchanan Mossel *Scribe: Miklos Racz*

In the previous lectures we discussed Bayesian martingale models; in these two lectures we turn our attention to non-martingale learning models. We first look at a Bayesian model introduced by Gale and Kariv [2], and then briefly discuss the special case of the Gale-Kariv model on the complete graph, as analyzed by Mossel and Tamuz [6]. Next, we study heuristic models of repeated voting: we first look at the DeGroot model, and then we analyze the repeated majority model.

1 Gale-Kariv model

We now study a model introduced by Gale and Kariv [2], which is more complicated than the models we have seen before. In this model there are two possible states of the world, and a finite number of agents trying to figure out the true state. In the beginning each agent receives an independent private signal picked from a distribution that depends on the state of the world, and from this they form a belief regarding the state of the world. Next, there are multiple rounds of voting in which each agent acts and observes simultaneously: they each make a declaration of the state of the world according to their current beliefs and observe the declarations of their neighbors. After each round, each agent calculates an updated posterior belief based upon what he or she observed. This is a Bayesian model, in the sense that the agents are Bayesian and their actions are aimed at maximizing their expected utility (i.e. figuring out the true state of the world) at each round.

1.1 Setup of the Gale-Kariv model

The state of the world is denoted by S, and it can take on two possible values: $+$ or $-$. A priori both states have the same probability, $1/2$.

Let $V = \{1, \dots, n\}$ denote the set of n agents, each agent trying to figure out the true state of the world. As suggested by the notation V, there is a graph structure in the background, the agents corresponding to vertices of this graph.

At the beginning, each agent receives an independent private signal giving some indication as to which is the true state of the world. More precisely: agent v receives a signal $s_v \in \Omega$ where $(\Omega, \mathcal{O})$ is a σ-algebra. What is the distribution of s_v? In the model we have two probability measures on $(\Omega, \mathcal{O})$, F^+ and F^-, and if the state of the world S is $+$, then s_v is distributed according to F^+, while if $S = -$, then s_v is distributed according to F^-.

We make some assumptions on F^+ and F^-. We assume that the Radon-Nykodim derivative $x(s) := \left.\frac{dF^+}{dF^-}\right|_s$ exists and is positive for all $s \in \Omega$. Furthermore, we assume that x is non-atomic and has a density over the reals. This is done to make the analysis simpler and to avoid situations in which an individual is undecided.

After receiving the private signal, each agent v forms a posterior belief: $\mathbb{E}(S|s_v)$, and based on this, $\operatorname{sgn} \mathbb{E}(S|s_v)$ is more likely to be the state of the world, according to the belief of v.

Example 1.1
Before moving on, let us look at a simple example. Suppose that $\Omega = [0,1]$, i.e. the signals are between 0 and 1, and suppose that F^+ and F^- both have densities, $f^+(s) = 2s$ and $f^-(s) = 2 - 2s$, respectively. In this case the Radon-Nykodim derivative / likelihood ratio is

$$x(s) = \left.\frac{dF^+}{dF^-}\right|_s = \frac{f^+(s)}{f^-(s)} = \frac{2s}{2-2s}.$$

If agent v receives a private signal s_v, the probability he or she assigns to the state of the world being $+$ is:

$$\mathbb{P}(S = +|s_v) = \frac{f^+(s_v)}{f^+(s_v) + f^-(s_v)} = \frac{2s_v}{2} = s_v.$$

So for instance if v receives the signal $s_v = 0.3$, then $\mathbb{E}(S|s_v = 0.3) = \mathbb{P}(S = +|s_v = 0.3) - \mathbb{P}(S = -|s_v = 0.3) = 0.3 - 0.7 = -0.4$ and so v—given no other information—believes that the state of the world is more likely to be $\operatorname{sgn}\mathbb{E}(S|s_v = 0.3) = -$, and there is also a "strength" to his/her belief given by: $\mathbb{P}(S = -|s_v = 0.3) = 0.7$.

Now let us turn back to describing the Gale-Kariv model. After receiving the private signals, there are multiple rounds of voting in which each agent acts and observes simultaneously. At time $t = 0$ (right after receiving the signals), the declaration of agent v is $d_v(0) := \operatorname{sgn}\mathbb{E}(S|s_v)$. Now each agent observes the declarations of his/her neighbors, and updates his/her belief of the state of the world accordingly. Then in each subsequent round the agents again make declarations and observe the declarations of their neighbors. The declaration of agent v at time t is

$$d_v(t) := \operatorname{sgn}\mathbb{E}(S|s_v, d_w(s) : w \in N(v), s < t),$$

where $N(v)$ denotes the neighbors of agent v in the underlying graph. After the declarations at time t, the belief of agent v is updated to $\mathbb{E}(S|s_v, d_w(s) : w \in N(v), s \le t)$.

One way to think about the process is the following. Suppose the agents are traders on the stock market, and the "state of the world" S being $+$ or $-$ indicates whether it is better to buy or sell in the market. The traders are busy and they don't have time to talk, the only way they communicate is when they pass each other in the hall / meet in the elevator, when they show each other a thumbs up or thumbs down, indicating their current beliefs of what the state of the world is (whether it is better to buy or sell).

Example 1.2
Let us again look at a simple example to understand how the process works. Suppose the setup (Ω, f^+, f^-) is the same as in Example 1.1, and suppose there are 3 agents, with the underlying graph being the triangle (i.e. the complete graph on 3 vertices). Suppose the received signals are $s_1 = 0.8, s_2 = 0.3, s_3 = 0.4$. The 3 agents calculate the likelihood ratio given their signal: $x(s_1) = 4, x(s_2) = 3/7, x(s_3) = 2/3$. Now the process starts, the players make their declarations. From the signals and likelihood ratios it is easy to see that the first round of declarations will be $d_1(0) = +, d_2(0) = -, d_3(0) = -$.

The agents then update their beliefs according to what they have seen. Let us look at how agent 1 updates his/her beliefs. Agent 1 sees that the other two agents declared $-$, so both of them received a signal between 0 and 1/2. Agent 1 then calculates the likelihood ratio of the things he has seen given that the real signal is $+$ and $-$, respectively. This likelihood ratio is the product of three likelihood ratios: the likelihood ratio of him receiving the signal $s_1 = 0.8$ and the likelihood ratios of agent 2 and 3 receiving a signal between 0 and 1/2. So the updated likelihood ratio of agent 1 is

$$\frac{f^+(0.8)}{f^-(0.8)} \left(\frac{\int_0^{1/2} f^+(s)\,ds}{\int_0^{1/2} f^-(s)\,ds} \right)^2 = \frac{1.6}{0.4} \left(\frac{1/4}{3/4} \right)^2 = \frac{4}{9}.$$

So the belief of agent 1 changes after the declarations, his/her new belief is

$$\mathbb{E}\left(S|s_1, d_2\left(0\right), d_3\left(0\right)\right) = \frac{4}{13} - \frac{9}{13} = -\frac{5}{13},$$

so in the next round agent 1 will declare $-$ as the state of the world. It can also be seen that the beliefs of agents 2 and 3 do not change after the initial declarations, so they too will declare $-$ as the state of the world in the next round. The agents thus reach consensus in the second round, and once consensus is reached, everything stays the same.

1.2 Filtrations and martingales in the Gale-Kariv model

There are many natural filtrations and consequently many natural martingales in the Gale-Kariv model. Let

$$\mathcal{F}_{v,t} := \sigma\left\{s_v, d_w\left(s\right) : w \in N\left(v\right), s < t\right\}$$

be the σ-algebra generated by the information that is available to agent v at time t and let

$$\mathcal{F}_v := \sigma\left\{\bigcup_{t\geq 0} \mathcal{F}_{v,t}\right\}$$

be the σ-algebra generated by the information that is available to agent v over the whole process. The belief of agent v at time t is $f_v\left(t\right) = \mathbb{E}\left(S|\mathcal{F}_{v,t}\right)$, which is a martingale with respect to the filtration $\mathcal{F}_{v,t}$ and which converges to $f_v := \mathbb{E}\left(S|\mathcal{F}_v\right)$ as $t \to \infty$.

Note that the declaration of agent v at time t is $d_v\left(t\right) = \operatorname{sgn} f_v\left(t\right)$. Define d_v to be $+$ if $f_v > 0$, to be $-$ if $f_v > 0$, and define it arbitrarily if $f_v = 0$. The convergence $\lim_t f_v\left(t\right) = f_v$ implies that if $f_v \neq 0$ then $\lim_t d_v\left(t\right) = d_v$. (We cannot say anything about the sequence of declarations if $f_v = 0$.) Furthermore, let $e_v = +$ if $\lim_t d_v\left(t\right) = +$, and otherwise let $e_v = -$ (this definition will be used in the next subsection).

Another natural filtration is the following. Let

$$\mathcal{G}_{v,t} := \sigma\left\{d_v\left(s\right) : s < t\right\}$$

be the σ-algebra generated by the information available to a friend of agent v at time t, who sees nothing else but the declarations of agent v at every round. As before, let

$$\mathcal{G}_v := \sigma\left\{\bigcup_{t\geq 0} \mathcal{G}_{v,t}\right\}$$

be the σ-algebra generated by the information that is available to the friend of agent v over the whole process. The belief of this friend at time t is $g_v\left(t\right) = \mathbb{E}\left(S|\mathcal{G}_{v,t}\right)$, which is a martingale with respect to the filtration $\mathcal{G}_{v,t}$ and which converges to $g_v := \mathbb{E}\left(S|\mathcal{G}_v\right)$ as $t \to \infty$.

Note: if $f_v > 0$ then $g_v > 0$, since "the friend of agent v knows that agent v is the professional". In fact more is true, the quality of prediction is the same for the friend and the agent. If the agent says only $+$ from one point on, then so will the friend. Similarly, if the agent says only $-$ from one point on, then so will the friend. If the declarations of the agent do not converge, then the agent has 1/2 - 1/2 chance of guessing the correct state of the world. But if the declarations of the agent do not converge, then the friend knows that the agent does not know the state of the world, and so the friend might as well flip a fair coin to decide.

1.3 Results and challenges in the Gale-Kariv model

The following result says that it is not possible that different agents converge to different actions. However, this does not rule out the possibility that the actions do not converge at all.

Proposition 1.1 (Gale, Kariv [2]). *Assuming that u is a neighbor of v then it cannot happen that $f_v > 0$ and $f_w < 0$.*

Proof. Our first claim is that the functions d_v and e_v both maximize the prediction probability of the signal given $\mathcal{F}_v$. Now let us introduce the filtration

$$\mathcal{G}_{v,w,t} := \sigma \left\{ d_v\left(s\right), d_w\left(s\right) : s < t \right\}$$

which is generated by the information available to a common friend of agents v and w and similarly as before let

$$\mathcal{G}_{v,w} := \sigma \left\{ \bigcup_{t \geq 0} \mathcal{G}_{v,w,t} \right\}.$$

Furthermore, let $c_{v,w}$ maximize the prediction probability given $\mathcal{G}_{v,w}$. Clearly

$$e_v \leq c_{v,w} \leq d_v,$$

since agent v has more information than the friend of agents v and w, who in turn has more information than the friend of agent v. But unless $f_v = 0$, $e_v = d_v$, and consequently $e_v = c_{v,w} = d_v$. Similarly, unless $f_w = 0$, $e_w = c_{v,w} = d_w$. So unless $f_v = 0$ or $f_w = 0$, we have $d_v = d_w$. This proves Proposition 1.1, since if we would have $f_v > 0$ and $f_w < 0$, then we would also have $d_v = +$ and $d_w = -$, which contradicts with the conclusion obtained in the previous sentence. □

In conclusion, we can say that the Gale-Kariv model is a very natural model to study, but there are many challenges related to it. For one, the model in its full generality is computationally intractable. Furthermore, it is very hard to tackle analytically, and consequently there are many open questions. It is not known how long it takes for the process to converge. It is also not known which networks aggregate well. It is provable that the complete graph and the star aggregate well (in the case of a star the central agent knows everything), and so do graphs with high degrees, e.g. graphs where every degree is at least of order $\log n$. However, it is not known whether trees or graphs with small diameter aggregate well.

1.4 Gale-Kariv model on the complete graph

Mossel and Tamuz consider the Gale-Kariv model on the complete graph in [6]. The discussion of this paper as seen on the slides was skipped during class, so we now just summarize the main results and refer the reader to the original article [6] for details. The main results are the following:

- Consensus is always reached, i.e. all agents converge to the same outcome a.s.
- The calculations of the agents are *efficient*. This is very important, in order for the framework to be implementable.
- Each round of voting improves the aggregation of information.
- The chance of a correct decision quickly approaches one as the number of agents increases. In particular, with high probability the agents reach consensus in the second round. (Proof idea: after the first vote, the individuals see n independent signals on the state of the world.)

2 Heuristic models

So far we haved discussed various Bayesian models for opinion updates. In this section we discuss simpler heuristic models of repeated voting. Why look at other models? First of all, real "agents" are probably not fully Bayesian, and instead apply some kind of heuristics (but which?). Also, a simpler update rule may lead to more complete analysis in understanding various features of the behavior.

In the following we analyze two models, the DeGroot model and the iterated majority model.

2.1 DeGroot model

The DeGroot model is the repeated weighted averaging of the opinions of one's self and one's neighbors. There are n individuals, denoted by $1, \ldots, n$, who at time $t = 0$ have some initial beliefs $f_0(i) = f(i)$ for $i \in [n]$. There are averaging weights $w_{ij} \geq 0$ which satisfy $\sum_{j=1}^{n} w_{ij} = 1$, and the update rule is $f_{t+1}(i) = \sum_{j=1}^{n} w_{ij} f_t(j)$. If $W = (w_{ij})_{i,j=1}^{n}$ denotes the stochastic matrix given by the weights and we think of the beliefs f_t at time t as a column vector, then the update rule in matrix notation is $f_{t+1} = W f_t$, so $f_t = W^t f$.

This linear model was introduced by Morris H. DeGroot (June 8, 1931 - November 2, 1989) in 1974 [1].

2.1.1 Examples

To familiarize ourselves with the model, let us look at a few examples.

- If $W = I$, i.e. $w_{ij} = \delta_{ij}$, then the network consists of n nodes and no edges, so nothing happens.
- If $w_{ij} = 1/n$ for all i, j, then everyone converges to the average of the initial beliefs in one round.
- Suppose $w_{ij} = 1/3$ for $(i - j) \mod n \in \{0, 1, n-1\}$, i.e. the graph consists of a cycle plus loops on the nodes. Again all beliefs will converge to the average of the initial beliefs. [Note that, by symmetry, if the beliefs converge, then all beliefs converge to the same belief.]
- Suppose $w_{ij} = 1/2$ for $(i - j) \mod n \in \{1, n-1\}$, i.e. the graph consists of a cycle on n nodes. The behavior in this case depends on the parity of n. If n is odd, then just like before, the beliefs converge to the average of the initial beliefs. However, if n is even then we might have a "blinking effect", see Figure 1.

$$\begin{array}{cc} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{array} \qquad\qquad \begin{array}{cc} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{array}$$

Figure 1: Blinking effect.

- Suppose the graph is a star with a loop at every node. That is, $w_{in} = w_{ii} = 1/2$ for $i < n$, and $w_{n,j} = 1/n$ for all j. In this case all beliefs converge to the same belief as before; however, this common belief will not be the average of all initial beliefs, but rather it will give greater weight to the initial belief of the central node of the star. The exact ratio can be calculated easily from the discussion below.

2.1.2 Markov chains and convergence

If μ_i denotes the row vector with 1 in the i^{th} coordinate and 0 elsewhere, then

$$f_t(i) = \mu_i f_t = \mu_i \left(W^t f\right) = \left(\mu_i W^t\right) f = \mu_{i,t} f = \mathbb{E}_{i,t}\left(f(X)\right),$$

where $\mu_{i,t}$ is the distribution of the Markov chain with transition matrix W started at i and run for t steps, and $\mathbb{E}_{i,t}$ denotes expectation according to this distribution. So we can see that the beliefs at time t are simply the expected beliefs according to the Markov chain with transition matrix W started at i and run for t steps. Due to this, we can use the theory of Markov chains to say things about the DeGroot model. In fact, almost everything we want to know follows from Markov chain theory. So in order to say things about the DeGroot model, we need to review some facts about Markov chains.

The total variation distance between two distributions P and Q (on a finite state space) is given by

$$d_{TV}(P, Q) = \frac{1}{2}\sum_x |P(x) - Q(x)|.$$

A finite state space Markov chain given by transition matrix P is called ergodic if there exists a finite k such that all entries of P^k are positive.

We know that if W corresponds to an ergodic chain, then there exists a unique stationary distribution π (i.e. $\pi W = \pi$). We also know that

$$d_{TV}(\mu_{i,t}, \pi) \to 0$$

as $t \to \infty$ for any i, i.e. no matter where the Markov chain starts, its distribution at time t converges to the stationary distribution (in the total variation metric) as $t \to \infty$. Furthermore, if $\mathbb{E}_\pi$ denotes expectation with respect to π, we also have

$$|\mathbb{E}_{i,t}(f(X)) - \mathbb{E}_\pi(f(X))| \to 0$$

as $t \to \infty$ for any i. In other words, if the chain is ergodic, then all beliefs converge to the same value, which is

$$\mathbb{E}_\pi(f(X)) = \sum_i \pi(i) f(i).$$

What is the stationary distribution π? If W is an ergodic random walk on an undirected graph, then the stationary distribution is proportional to the degrees of the nodes, i.e.

$$\pi(i) = \frac{\deg(i)}{\sum_j \deg(j)}.$$

So the importance of a node is determined by the number of neighbors it has.

2.1.3 Examples revisited

Let us look at the examples from Section 2.1.1 again.

- If $W = I$, i.e. $w_{ij} = \delta_{ij}$, then the network consists of n nodes and no edges, so nothing happens. The Markov chain corresponding to W is evidently not ergodic.
- If $w_{ij} = 1/n$ for all i, j, then the corresponding Markov chain is ergodic, with uniform stationary distribution. Furthermore, the Markov chain reaches stationarity in one round, so everyone converges to the average of the initial beliefs in one round.
- Suppose $w_{ij} = 1/3$ for $(i-j) \mod n \in \{0, 1, n-1\}$, i.e. the graph consists of a cycle plus loops on the nodes. The corresponding Markov chain is ergodic with uniform stationary distribution, so all beliefs will converge to the average of the initial beliefs.
- Suppose $w_{ij} = 1/2$ for $(i-j) \mod n \in \{1, n-1\}$, i.e. the graph consists of a cycle on n nodes. The behavior in this case depends on the parity of n. If n is odd, then the corresponding Markov chain is ergodic with uniform stationary distribution, so just like before, the beliefs converge to the average of the initial beliefs. However, if n is even, then the chain is not ergodic, and consequently we might have a "blinking effect", as shown above.
- Suppose the graph is a star with a loop at every node. That is, $w_{in} = w_{ii} = 1/2$ for $i < n$, and $w_{n,j} = 1/n$ for all j. The corresponding Markov chain is ergodic, so all beliefs converge to the same belief as before; however, this common belief will not be the average of all initial beliefs, since the stationary distribution is not the uniform distribution. The common belief will give greater weight to the initial belief of the central node of the star, since the central node has a higher weight in stationarity. Calculation shows that $\pi_i = \frac{2}{3n-2}$ for $i < n$ and $\pi_n = \frac{n}{3n-2}$, so the common belief will be

$$\mathbb{E}_\pi (f(X)) = \sum_i \pi(i) f(i) = \frac{2}{3n-2} \sum_{i=1}^{n-1} f(i) + \frac{n}{3n-2} f(n).$$

2.1.4 Rate of convergence

An important question is the rate of convergence to the common opinion. In the theory of Markov chains there are many ways to measure the rate of convergence of a chain. We can directly use any of them to obtain bounds on the rate of convergence to the common opinion via

$$\max_i |f_t(i) - \mathbb{E}_\pi (f(X))| \leq 2 \|f\|_\infty \max_i d_{TV}(\mu_{i,t}, \pi),$$

where $\|f\|_\infty = \max_i f(i)$.

The techniques include spectral gap bounds, conductance bounds, and many more (log-Sobolev bounds, coupling bounds, etc.). We now briefly talk about these.

Spectral gap bounds. Suppose our ergodic Markov chain has transition matrix P, and let g be the smallest non-zero eigenvalue of $I - (P + P^*)/2$, where I is the identity matrix. This is known as the spectral gap. Then for

$$t = s \frac{1 + \max_i \ln(1/\pi_i)}{g}$$

we have

$$\max_i d_{TV}(\mu_{i,t}, \pi) \leq e^{-s}.$$

Conductance bounds. Define the edge measure Q as

$$Q(x,y) := w_{x,y}\pi(x) + w_{y,x}\pi(y) \qquad\qquad Q(A,B) := \sum_{x \in A, y \in B} Q(x,y),$$

and define the conductance of the Markov chain as

$$\ell = \min_{\pi(A)\leq\frac{1}{2}} \frac{Q\left(A,A^c\right)}{\pi\left(A\right)}, \tag{1}$$

which is also known as the bottleneck ratio of the chain. Then Cheeger's inequality says

$$\frac{\ell^2}{8} \leq g,$$

and so for time

$$t = 8s\frac{1+\max_i \ln\left(1/\pi_i\right)}{\ell^2}$$

we have

$$\max_i d_{TV}\left(\mu_{i,t},\pi\right) \leq e^{-s}.$$

This means that if there are no "isolated communities" or "bottlenecks" in the graph, then convergence is quick. Note: in the minimum in (1) it is enough to look at connected sets.

Exercise. Compute the spectral gap of each of the Markov chains considered in Section 2.1.1.

Exercise. Compute the spectral gap of the transition matrix P given by a random walk on the graph G which consists of two K_n components connected with one single edge. This single edge is a "bottleneck" in the graph, so the Markov chain mixes slowly.

2.1.5 Cheaters in the DeGroot model

What if somebody cheats? Can they convince the rest of the group to reach whatever value they want?

Proposition 2.1. *If the chain is ergodic and there is one cheater by repeatedly stating a fixed value, then all opinions will converge to that value.*

Proof. First, assume $w_{ij} > 0$ for all $i,j \in [n]$. Let player 1 be the cheater, i.e. $f_t\left(1\right) = x$ for all $t \geq 0$. We then know that

$$f_t\left(i\right) = w_{i1}x + \sum_{j=2}^{n} w_{ij} f_{t-1}\left(j\right). \tag{2}$$

Now the map $f_t \to f_{t+1} = Wf_t$ is a contraction with contraction constant $c = 1 - \min_i w_{i1} < 1$ since by (2) we have

$$\begin{aligned}
\sup_{i\in[n]} \left|f_t\left(i\right) - g_t\left(i\right)\right| &= \sup_{i\in[n]} \left|\sum_{j=2}^{n} w_{ij}\left(f_t\left(j\right) - g_t\left(j\right)\right)\right| \\
&\leq \sup_{i\in[n]} \sum_{j=2}^{n} w_{ij}\left|f_t\left(j\right) - g_t\left(j\right)\right| \\
&\leq c \sup_{j\in[n]} \left|f_t\left(j\right) - g_t\left(j\right)\right|.
\end{aligned}$$

From (2) it is easy to see that the vector containing x in every coordinate is a fixed point of the map. By the Banach fixed point theorem this is the unique fixed point of the map, and furthermore f_t converges to this fixed point as $t \to \infty$.

If not all w_{ij} are positive, then ergodicity of the chain guarantees that there exists a finite k such that W^k contains strictly positive entries. We can then look at the map at every k steps, this new map will be a contraction, and the proof follows as above. The details are left to the reader. $\square$

Proposition 2.2. *Suppose the chain is ergodic and m out of the n players are cheaters. Then all players' declarations converge, and the final signal only depends on the cheaters' declarations.*

Proof. Let us number the nodes from 1 to n, numbering the cheaters from $n-m+1$ to n. Write the stochastic matrix W in the following block matrix form:

$$W = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where A is an $(n-m)\times(n-m)$ matrix, B is an $(n-m)\times m$ matrix, C is an $m\times(n-m)$ matrix, and finally D is an $m\times m$ matrix. Now

$$Q = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix}$$

is the transition matrix corresponding to the case when there are m cheaters, i.e. if there are m cheaters then $f_{t+1} = Qf_t$, and so $f_t = Q^t f$. Now calculation shows that

$$Q^t = \begin{bmatrix} A^t & \left(\sum_{k=0}^{t} A^k\right) B \\ 0 & I \end{bmatrix}.$$

We know that Q corresponds to a Markov chain which has the nodes of the cheaters as absorbing states. The matrix A^t gives the probabilities that we started the Markov chain at a non-cheater node and at time t we are still at a non-cheater node. Assuming the graph is connected, the Markov chain governed by Q gets absorbed eventually, so $A^t \to 0$ as $t\to\infty$. Consequently

$$Q^t \to Q_\infty := \begin{bmatrix} 0 & (I-A)^{-1}B \\ 0 & I \end{bmatrix}$$

as $t\to\infty$, and so

$$f_t = Q^t f \to f_\infty := Q_\infty f.$$

Since the first $(n-m)$ columns of Q_∞ are 0, f_∞ only depends on the cheaters' declarations. If we write x for the column vector of length m containing the cheaters' declarations, then we have

$$f_\infty = \begin{bmatrix} (I-A)^{-1}B \\ I \end{bmatrix} x.$$

□

2.2 Iterated majority

A model that avoids this "cheating effect" is repeated majority. The setup is the following. Now the original signals $s_i(0)$ are $+$ or $-$. There are again weights w_{ij} (this time they do not have to be normalized) and the update rule is

$$s_{t+1}(i) = \operatorname{sign}\left(\sum_{j=2}^{n} w_{ij} s_t(j)\right).$$

This model is a non-linear analogue of the DeGroot model. In the following we assume that the weights are chosen in such a way that ties (i.e. when the weighted sum of signals is 0) are impossible. A way of achieving this is the following: nodes of odd degrees take the majority of their neighbors, while nodes of even degrees take the majority of their neighbors and themselves. [Another solution would be to toss a fair coin in case of a tie.]

Cheating in this model is hard, consider the following two examples.

- Suppose the graph of the network of is a star, and the number of $+$'s initially is at least 3 greater than the number of $-$'s initially. Then a leaf cheating and saying $-$ in every round does not change the other peoples' final opinions (which will be $+$).
- Similarly, consider a complete graph, where the number of nodes with a $+$ signal initially is at least 4 greater than the number of nodes with a $-$ signal initially. Then a cheating node that says $-$ in each round does not change the other peoples' final opinions (which will be $+$).

To familiarize ourselves with the model, let us look at a few examples.

- If $W = I$, i.e. $w_{ij} = \delta_{ij}$, then the network consists of n nodes and no edges, so nothing happens.
- If $w_{ij} = 1$ for all i, j and n is odd, then this is exactly simple majority.
- Suppose $w_{ij} = 1$ for $(i - j) \mod n \in \{0, 1, n - 1\}$, i.e. the graph consists of a cycle plus loops on the nodes. Then there is the possibility of a "blinking effect", see Figure 2. In the same example, there

$$\begin{array}{cc} - & + \\ + & - \end{array} \qquad\qquad \begin{array}{cc} + & - \\ - & + \end{array}$$

Figure 2: Blinking effect.

are fixed points which are not all $+$ or all $-$, see Figure 3.

$$\begin{array}{ccccccc} & - & - & - & - & - & \\ - & & & & & & + \\ & + & + & + & + & + & \end{array}$$

Figure 3: Fixed point which is not all $+$ nor all $-$.

- Suppose the graph is a star with a loop in the middle. That is, $w_{ij} = 0$ if $i, j < n$, and $w_{i,j} = 1$ if $\max\{i, j\} = n$. Suppose further that n is odd. Then after the first round the central node changes to simple majority, and in the second round each leaf copies this vote from the central node. So there is convergence to the majority vote in two rounds.

2.2.1 Convergence

In general we can ask: does the iterated majority process converge? If so, to what? If it does not converge, what can we say about the dynamics? Many of these questions are open!

From the example in Figure 2 we see that the process does not necessarily converge. In general, if the graph is a bipartite graph and at time $t = 0$ one component has all $+$ votes, the other component has all $-$ votes, then we see the same blinking effect among the two components. More generally, suppose we can partition the graph into two parts such that each node has more neighbors in the component they do not belong to. Then if at time $t = 0$ one component has all $+$ votes and the other component has all $-$ votes, then we see the same blinking effect among the two components.

Since there are at most 2^n possible declarations of the n nodes as a whole, we know that if there is no convergence then the declarations will fall into a cycle. A trivial upper bound for the length of the cycle is 2^n. In case of directed graphs examples can be created where the length of the cycle is exponential in n. What is a tight upper bound for the length of the cycle for directed graphs? How about for undirected graphs: what is the longest possible cycle? These are open questions.

Even if the process converges, it is hard to say what it converges to from an arbitrary starting condition.

2.2.2 Convergence for biased signals

What happens for biased signals? Some partial results are known for asynchronous dynamics. In this setup the updates do not all occur at the same time, but rather there is an exponential clock (corresponding to a Poisson process) at every node, and when this rings the node updates its signal to the majority of its neighbors (flipping a fair coin in case of a tie). The following results come from statistical physics. For infinite grids in dimensions 2 and higher ($\mathbb{Z}^2$ is the "social network of beggars in New York"), and for regular trees there exists a $q < 1$ such that if the signals are iid with $\mathbb{P}(s = +) = p > q$ then almost surely all signals converge to $+$. Howard studied the process on infinite 3-regular trees in [4], showing that there exists a $p > 1/2$ for which there is no convergence to all $+$ signals. So combining the two results we have the following picture for infinite 3-regular trees: there exist $1/2 < q_1 \leq q_2 < 1$ such that for $1/2 \leq p \leq q_1$ there is no convergence to all $+$ signals, for $p > q_2$ there is convergence to all $+$ signals, and for $q_1 \leq p \leq q_2$ we do not know what happens.

The process on finite and infinite graphs behaves very differently, which can be seen on the following example.

Proposition 2.3. *Let T_n be the n-level binary tree (with a loop at the root to make the root's degree odd), and let $1/2 < p < 1$. Then the probability of convergence to all $+$ goes to 0 as $n \to \infty$.*

Proof. If there is a cherry at the bottom (a node at level $n - 1$ and its two children) where all three nodes have a $-$ signal, then these signals will stay like this forever. The probability of no cherry at the bottom having all $-$ signals goes to 0 as $n \to \infty$. $\square$

An interesting question is the following problem (which may or may not be open): find a big family of graphs with only all $+$ or all $-$ as fixed points. The previous paragraph shows that binary trees do not provide such an example. On the other hand, complete graphs do. Are there any other examples?

What is an example of families of graphs for which majority dynamics aggregates well? Expander graphs provide such an example. A graph $G = (V, E)$ is an (e, s) expander if for every $S \subseteq V$ such that $|S| \leq s\,|V|$ we have

$$|\{v \in S^c : \exists\; u \in S \text{ such that } (u, v) \in E\}| \geq e\,|S|\,.$$

The following statement (which we do not prove, but which can be found in the survey by Hoory, Linial, and Wigderson [3]) shows that expanders exist, and that actually there are many of them.

Proposition 2.4. *There exists $a > 0$ such that for large k and n, a random k-regular graph G on n vertices is a $(0.75k, a/k)$ expander with high probability.*

The next result gives an answer to the question asked at the beginning of the paragraph.

Proposition 2.5. *For k-regular $(0.75k, a/k)$ expanders on n vertices (with k and n large enough), if $p > 1 - a/(2k)$ then the graph aggregates well with high probability. In fact, we show that if at time $t = 0$ a fraction of at least $1 - a/(2k)$ of the voters receive a $+$ signal, then the dynamics will converge to all $+$ signals.*

Proof. The proof is from Kanoria and Montanari [5]. We show that the number of nodes with a $-$ signal contracts by 5/6 at each iteration. Let S denote the set of vertices with a $-$ signal at time $t = 0$, let S^c denote the set of all other vertices, and let $s := |S|$. Define the following quantities:

$$\begin{aligned} n(+) &:= |\{v \in S^c : v \text{ has at least 1 but at most } k/2 \text{ neighbors in } S\}| \\ n(-) &:= |\{v \in S^c : v \text{ has at least } k/2 \text{ neighbors in } S\}| \\ l &:= \#\text{ of edges from } S \text{ to itself.} \end{aligned}$$

Since $s \leq \frac{a}{2k} n$ by the hypothesis, we can apply the expansion condition to S, which says that

$$n(+) + n(-) \geq 0.75ks. \tag{3}$$

Now let us count the sum of the degrees of vertices in S. On the one hand, this quantity is exactly ks since our graph is k-regular. On the other hand, there are l edges connecting two nodes in S, each edge contributing 2 to the sum; there are $n(+)$ vertices in S^c contributing at least 1 to the sum; and there are $n(-)$ vertices in S^c contributing at least $k/2$ to the sum. This gives us

$$ks \geq 2l + n(+) + \frac{k}{2} n(-). \tag{4}$$

Summing (3) and (4) we arrive at

$$0.25ks \geq 2l + \left(\frac{k}{2} - 1\right) n(-),$$

dividing by $k/2$ gives us

$$0.5s \geq \frac{2l}{k/2} + \left(1 - \frac{2}{k}\right) n(-),$$

from which it follows that

$$\frac{2l}{k/2} + n(-) \leq \frac{0.5}{1 - 2/k} s \leq \frac{5}{6} s, \tag{5}$$

using that $k \geq 5$. We know that the number of vertices with a $-$ signal after the first iteration is at most $2l/(k/2) + n(-)$; there are $n(-)$ vertices that had a $+$ signal and now changed to a $-$ signal, and there are at most $2l/(k/2)$ vertices in S which continue to have a $-$ signal after this iteration. So (5) says that S is a contracting set. ☐

Concluding remarks. We have seen many examples where it is easy to establish that not all voters converge to the same vote. In some rare examples it is possible to show convergence to the same vote, using strong bias and expansion arguments.

2.2.3 Asking a different question

So far we analyzed whether or not repeated majority results in achieving consensus. Since we know that in real life consensus is often not reached, we should perhaps aim for less: what is the effect of repeated majority on the outcome of the vote?

Let us look at the following setup. Originally voters receive a $+$ or $-$ signal which is correct with probability $p > 1/2$. We know that the optimal fair vote is to take the majority of the original signals. What we do instead is let the voters learn, converse among each other, and change their opinions for d days. This is done by each day updating their view to the majority of the people they appreciate. Then at the end of the d days we take a majority vote. Does this make the quality of the vote worse? We have seen in the electoral college example that it destroys optimality. But does it destroy aggregation? The following example shows that it can.

Example 2.1 ***US Media***
Suppose there are 100 media outlets and 2000000 voters. Each media outlet takes a majority vote among some number of other media outlets, while each voter takes a majority among 1-10 friends and 30-50 media outlets. Does aggregation hold if we repeat for d days? No! It is enough that the media gets the wrong signal in the first round.

This is perhaps not such a realistic example, but it is possible to construct more realisic ones. In particular it suffices that a positive fraction of the population gives more weight to the media than their friends.

Example 2.2 ***US Media, part 2***
Suppose 10% of the people give more weight to the media than to their friends. Suppose 10% of the people give more weight to the media and people in the previous group than to the rest of their friends. And so on, for say 60% of the voters. This does not aggregate either.

In the previous two examples we had directed graphs: voters took into account media outlets, but media outlets did not take into account voters. Can we construct such an example? Yes, see the following.

Example 2.3 ***US Media, part 3***
Suppose there is 1 media outlet, and the n voters are partitioned into pairs. The media outlet takes majority over all voters, while voters take the majority of the view of the outlet, the view of their pair, and the view of their own self. Suppose the true signal is given with probability $p = 0.51$. Now if initially the media outlet gives the wrong signal, then after the first round a fraction of roughly $1 - 0.51^2$ of the pairs will have the wrong signal, and these pairs will remain so for all following rounds.

We see that repeated majority can kill the aggregation of information. Are there any situations where it does not? Yes, as usual the democratic examples are good. For instance, the following are good examples.

- Suppose the voters are partitioned into many groups and in each group we do repeated majority among all members of the group. Each group decision is a monotone fair function, so it is correct with probability at least p. If there are many groups, then by the law of large numbers we have aggregation.
- Suppose the voters are partitioned into 5 groups, each of size $n/5$. Then in each group we reach the correct decision with probability tending to 1, so there is aggregation of information.

Are there more interesting examples? Let G be the directed graph where $u \to v$ means that u takes v as part of its majority rule. We say that G is (vertex-)transitive if there exists a group Γ acting transitively on $V(G)$ mapping edges to edges.

Proposition 2.6 (Mossel-Tamuz)**.** *If G is transitive then we have aggregation of information.*

Proof. The overall aggregation function is fair, monotone, and transitive! $\square$

In fact a much weaker condition suffices.

Proposition 2.7 (Mossel-Tamuz)**.** *Let m be the minimal size of an orbit under the action of Γ. Then we have aggregation of information at $p = 1/2 + 1/\log m$.*

Moral: it is fine to have newspapers as long as there are many!

References

[1] M.H. DeGroot. Reaching a Consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[2] D. Gale and S. Kariv. Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346, 2003.

[3] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[4] C.D. Howard. Zero-temperature Ising spin dynamics on the homogeneous tree of degree three. *Journal of Applied Probability*, 37(3):736–747, 2000.

[5] Y. Kanoria and A. Montanari. Majority dynamics on trees and the dynamic cavity method. *Arxiv preprint arXiv:0907.0449*, 2009.

[6] E. Mossel and O. Tamuz. Making Consensus Tractable. *Arxiv preprint arXiv:1007.0959*, 2010.