Evolution of Information Gathering Strategies

Eleanor Brush
October 31, 2014

Contents

1 Abstract

Birds in flocks, like animals in many types of social groups, can use their peers to learn about the environment and update their opinions about where to move in that environment. Previous work on starlings has measured the number of other birds an individuals pays attention to and shown that this number of "neighbors" leads to a social network that is conducive to the whole flock reaching consensus. It is unclear, however, why an individual would change its behavior, i.e. the number of neighbors it learns from, to achieve this outcome. Individual birds should optimize how well they learn about the environment. There are (at least) two features of the environment that are important to an individual bird's fitness—the location of a predator and the location of resources—and it can try to learn about these from its peers. In this work, we identify the optimal strategies for birds under selection pressure dictated by these types of information and explore the relationship between the robustness measure and the correlation length of the flock, a measure commonly used to describe how close a flock's structure is to criticality.

2 Model

2.1 Learning dynamics An individual's opinion changes with a probability that depends on its social interactions and the strength of the environmental stimulus. Specifically, an individual's opinion σ_i can be either 1 or -1 and σ_i switches with probability $w_i(\sigma_i) = \frac{1}{2}(-\sigma_i \sum_j A_{ij}\sigma_j)(1-\beta_i\sigma_i)$ where A_{ij} indicates the strength of the connection from j to i and β_i indicates the strength of the environmental stimulus perceived by i. The stochastic dynamics of the opinions can be written down with a master

equation [4]. The vector $\vec{v}(t)$ of expected values, $q_i(t) = \langle\sigma_i(t)\rangle$, satisfies the following differential equations: [4]

$$\frac{d\vec{v}}{dt} = M\vec{v} + \vec{\beta}^T(\vec{1} - \vec{v}).$$

We simplify this by writing $\frac{d\vec{v}}{dt} = A\vec{v} + \vec{\beta}^T\vec{1}$, where $A = M - B$ and B is a diagonal matrix with the values of $\vec{\beta}$ along the diagonal. In the absence of the environmental signal, there is no trend in the dynamics. In the presence of the environmental signal, however, those individuals who perceive the signal will be more likely to switch to having an opinion of 1 so their expected value will increase, and consequently so will the values of the those paying attention to the informed individuals. Thus, having a higher expected value can be interpreted as having more information about the information and having a lower expected value can be interpreted as having less information.

2.2 Individual-level fitness To find the optimal strategies from the individuals' perspectives, we need a measure of individual-level fitness, which we define as follows. Each individual has a strategy that determines how many other individuals he pays attention to, $n_i = \sum_{j\neq i}\mathbf{I}(A_{ij} \neq 0)$. We distribute the individuals randomly in space. Each individual then pays attention to its n_i nearest neighbors. One individual is chosen to be a receiver; it and all those individuals within a radius r of the receiver perceive an external signal of strength β. We find the expected opinion of each individual q_i after a period of time T. We refer to each such set of random positions and receiver as a signaling event.

We consider two types of signals: predators and food. If the signal is information that a predator is present, we assume that whichever individual is least aware of the signal will be predated, i.e. the individual i such that q_i is lowest. We therefore find, over many signaling events, the probability that an individual will be eaten and an individual's fitness is given by the probability of surviving, i.e. 1 minus the probability of being predated. If the signal is information that resources are available, we assume that whichever individual is most aware of the signal will get access to the resource, i.e. the individual i such that q_i is highest. In this case, fitness is proportional to the probability of being the first to reach the resources.

2.3 Group-level performance We consider two group-level properties that indicate how cohesive the group is. The first property is $\mathscr{H}_2$ robustness, as used in [9, 8]. This is a measure of the robustness of the consensus state in which all birds have the same opinion to noise at equilibrium, as used in (see Appendix Sec. 6.1). The second property is the correlation length of the flock. This is the distance at which the average correlations between birds changes from positive to negative, i.e. the distance over which birds' opinions tend to be positively correlated with each other (see Appendix Sec 6.2).

2.4 Optimization methods To understand what strategies we might expect to find, we are interested in identifying the optimal strategies. We does this in two ways. First, using the framework of adaptive dynamics in finite populations, we identify the evolutionarily stable strategy. This framework assumes that there is a homogeneous population into which a mutant individual tries to invade. If strategies change on a learning rather than an evolutionary timescale, or if there is just a lot of variation in the population, we might expect individuals to try to optimize their strategies in the context of a heterogeneous population. This is our second method of optimization. Given a random set of initial strategies over the group, we allow each bird to choose the strategy that would be best given the rest of the strategies being used and repeat this process until the birds reach an equilibrium set of strategies. We repeat this over many initial sets of strategies to find average properties of this optimization process.

3 Results

3.1 ESS strategies Surprisingly, even without imposing costs on paying attention to more neighbors, paying attention to the whole flock is not usually the best strategy. The ESS strategy of the number of neighbors to pay attention strongly depends on which selection pressure is being applied. If a focal

bird pays attention to many others, its opinion will be similar to theirs. This will prevent it from being the most informed and hence the first to locate resources. Thus, the uninvadable strategy is to pay attention to few neighbors when selection is due to knowledge about resources. On the other hand, having many neighbors will prevent the focal bird from being the least informed and thus the most likely to be eaten. Thus, the uninvadable strategy is to pay attention to many neighbors when selection is due to knowledge about predators. Figure 1 shows the ESS strategy when selection is due to knowledge about resources is always lower than the ESS strategy when selection is due to knowledge about predators (Figure 1).

The ESS strategies when selection is due to knowledge about predators show two interesting features. First, there are many parameters for which there are multiple ESS strategies. It might seem that if a strategy outperforms all others when it is the resident strategy, it should also outperform other strategies as an invader. However, when there are multiple ESS strategies, if we take any two of those ESS strategies, each is uninvadable by the other. This indicates how important the social environment set by the resident strategy is. The resident strategy essentially sets a level of background noise. As an example, consider the strategies of having sixteen or seventeen neighbors when the radius of the signal is $r = 0.1$. Figure 1 shows how the probability of being predated for each of these invasion scenarios depends on the number of birds in the group and the number of informed neighbors the invader has. A bird with seventeen neighbors will have more uninformed neighbors than one with sixteen neighbors and thus will have less information itself. This poor information is costly when the resident strategy is to have sixteen neighbors since the residents will have less noisy information, which prevents an mutant with seventeen neighbors to invade. Figure 1 shows that by paying attention to more uninformed neighbors this mutant is likely to be eaten when there are many truly informed birds. However, having less noisy information does not allow a sixteen neighbor mutant to invade a seventeen neighbor group since, regardless of the number of uninformed neighbors the invader has, the background level of noise is much higher in a group all of whom have more uninformed neighbors. The sixteen neighbor mutant also pays a cost by being less likely to pay attention to the right neighbors when there are very few truly informed birds. Figure 1 shows that by being less likely to pay attention to the right neighbors this mutant is likely to be eaten when there are few truly informed birds.

These opposing forces caused by the costs of paying attention to more uninformed neighbors and the costs of not paying attention to informed neighbors also explains the second feature of the ESS strategies: they are non-monotonic with respect to the the radius of the signal, i.e. how public the signal is (Figure 1).

3.2   Learned strategies The equilibrium strategies when the individuals are allowed to learn are similar to the ESS strategies: learned strategies when selection is due to predation are always higher than learned strategies when selection is due to resources (Figure 2). Surprisingly, even when the group starts with heterogeneous strategies, they often reach a homogeneous equilibrium. Under selection due to predation, the $\mathscr{H}_2$ robustness and correlation length of the flock tend to decrease, whereas under selection due to resources, $\mathscr{H}_2$ robustness is relatively constant and correlation length tends to increase (Figure 2).

3.3   Relationship between group properties Figure 2 shows that there is a relationship between two group properties: the $\mathscr{H}_2$ norm and the correlation length of the flock.

4   Discussion

explore vs. exploit risk aversion vs. risk tolerance
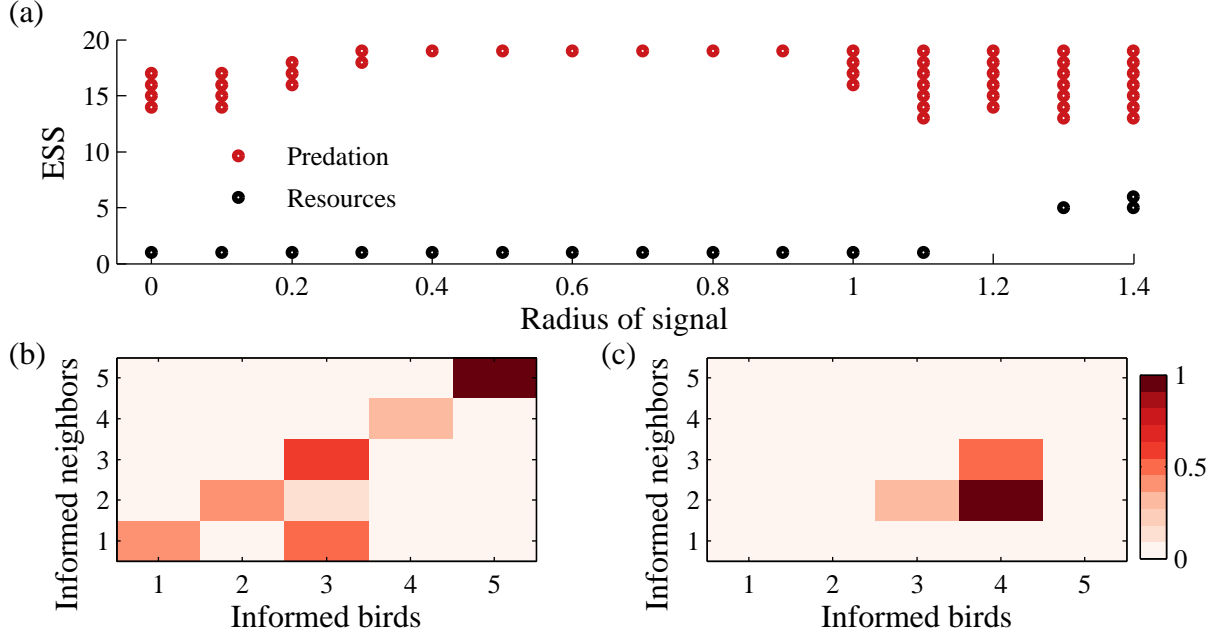
5   Figures

Figure 1: The ESS number of neighbors is always higher if selection is due to predation than due to resources. When selection is due to predation, there can be multiple ESS strategies, which are a non-monotonic function of the radius of the signal. In (a), we show the (possibly multiple) ESS strategies as a function of the radius of the signal. In (b) and (c), each heat map shows the probability that an invader is eaten as a function of the number of informed birds in the flock on the horizontal axis and the number of informed neighbors the invader has on the vertical axis. In (b), the resident strategy is to have 16 neighbors and the invader strategy is to have 17 neighbors. In (c), the resident strategy is to have 17 neighbors and the invader strategy is to have 16 neighbors. Parameters: in all panels $\beta = 1$, $T = 1$, in (b) and (c) $r = 0.1$.

6 Appendix

6.1 H2 Norm Let y represent the distance between the vector of opinions and consensus, i.e. $y = QxQ^T$, where $Q \in \mathbb{R}^{n-1 \times n}$ is such that $Q\vec{1} = 0$, $QQ^T = I_{n-1}$, and $Q^T Q = I_n - \frac{1}{n}\vec{1} \times \vec{1}^T$. Then

$$\dot{y}(t) = -\overline{L}y(t) + Q\xi(t)$$

where $\overline{L} = QLQ^T$. As shown in [9], if $\Sigma_y(t) = \mathbb{E}[y(t)y(t)^T]$, Σ_y at equilibrium solves

$$0 = -\overline{L}\Sigma_y - \Sigma_y \overline{L}^T + D. \tag{1}$$

$$\mathscr{H}_2 = \sqrt{\mathrm{Tr}(\Sigma_y)}$$

$$\mathscr{H}_2 \text{ robustness} = \frac{1}{\mathscr{H}_2}$$

6.2 Calculating correlations The following is derived from the work in [2].

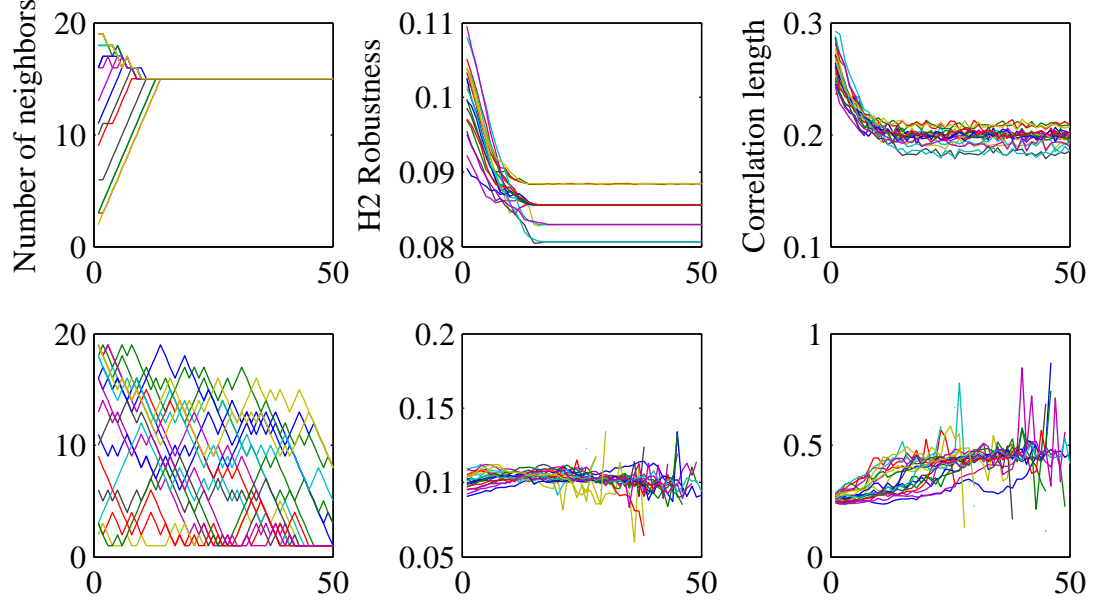$$H(\vec{v}) = -\vec{v}^T A\vec{v} - \vec{\beta}^T \vec{v}$$

Figure 2: When the birds choose the optimal number of neighbors, given the strategies the rest of the flock are using, they settle on higher strategies when selection is due to predation than when selection is due to resources. Under selection due to predation, the $\mathcal{H}_2$ robustness and correlation length of the flock tend to decrease, whereas under selection due to resources, $\mathcal{H}_2$ robustness is relatively constant and correlation length tends to increase. The upper row shows results from implementing selection based to predation and the lower row shows results from implementing selection based on resources. In each panel, the x-axis represents the number of times the birds are allowed to choose optimal strategies. The first column shows one example of how the birds' strategies change over time. The second column shows, for many initial conditions, how the $\mathcal{H}_2$ robustness changes over time. The third column shows, for many initial conditions, how the correlation length changes over time. Parameters: $\beta = 1$, $r = .1$, $T = 1$.

Define $V = \sum_k v_k / N$ and $\vec{y} = (\vec{v} - V\vec{1})/V$. Then $\vec{v} = V(\vec{1} + \vec{y})$.

$$H(\vec{y}, V) = -V^2(\vec{1}^T + \vec{y}^T)A(\vec{1} + \vec{y}) - \vec{\beta}^T V(\vec{1} + \vec{y})$$
$$= -V^2(\vec{1}^T + \vec{y}^T)A\vec{y} - V\vec{\beta}^T\vec{y} - V\vec{\beta}^T\vec{1} \text{ since } A\vec{1} = \vec{0}$$
$$= -V^2\vec{y}^T A\vec{y} + (-V^2\vec{1}^T A - V\vec{\beta}^T)\vec{y} - V\vec{\beta}^T\vec{1}$$

Define $Q \in M_{N-1,N}$ that rotates an N-vector away from the consensus vector and define $\vec{z} = Q\vec{y}$ so that $\vec{y} = Q^T\vec{z}$ since $\sum_k y_k = 0$.

$$H(\vec{z}, V) = -V^2\vec{z}^T QAQ^T\vec{z} + (-V^2\vec{1}^T AQ^T - V\vec{\beta}^T Q^T)\vec{z} - V\vec{\beta}^T\vec{1}$$
$$= V^2\vec{z}^T P\vec{z} + (-V^2\vec{\sigma}_1^T - V\vec{\sigma}_2^T)\vec{z} - V\vec{\beta}^T\vec{1}$$
$$\Rightarrow P(\vec{z}, V) = \frac{1}{\mathscr{Z}}\exp\left(-V^2\vec{z}^T P\vec{z} + (V^2\vec{\sigma}_1^T + V\vec{\sigma}_2^T)\vec{z} + V\vec{\beta}^T\vec{1}\right)$$

where $P = -QAQ^T$, $\vec{\sigma}_1 = QA^T\vec{1}$, and $\vec{\sigma}_2 = Q\vec{\beta}$.

Fact 1

$$\int \exp\left(-\frac{1}{2}\vec{z}^T P \vec{z} + \vec{\sigma}^T \vec{z}\right) d^{N-1}z = \sqrt{\frac{(2\pi)^{N-1}}{\det P}} \exp\left(-\frac{1}{2}\vec{\sigma}^T P^{-1}\vec{\sigma}\right)$$

$$\Rightarrow \int \exp\left(-\vec{z}^T P \vec{z} + \vec{\sigma}^T \vec{z}\right) d^{N-1}z = \sqrt{\frac{(2\pi)^{N-1}}{2^{N-1}\det P}} \exp\left(-\frac{1}{4}\vec{\sigma}^T P^{-1}\vec{\sigma}\right)$$

$$= \sqrt{\pi^{N-1}}\sqrt{\frac{1}{\det P}} \exp\left(-\frac{1}{4}\vec{\sigma}^T P^{-1}\vec{\sigma}\right)$$

$$\mathscr{Z} = \int_{-\infty}^{\infty} \int_{\mathbb{R}^{N-1}} \exp\left(V^2\vec{z}^T P \vec{z} + (V^2\vec{\sigma}_1^T + V\vec{\sigma}_2^T)\vec{z} + V\vec{\beta}^T \vec{1}\right) d^N z\, dV$$

$$= \int_{-\infty}^{\infty} \exp(V\vec{\beta}^T\vec{1}) \int_{\mathbb{R}^{N-1}} \exp\left(V^2\vec{z}^T P \vec{z} + (V^2\vec{\sigma}_1^T + V\vec{\sigma}_2^T)\vec{z}\right) d^N z\, dV$$

$$= \sqrt{\pi^{N-1}} \int_{-\infty}^{\infty} \exp(V\vec{\beta}^T\vec{1}) \times \left(\sqrt{\frac{1}{V^{2(N-1)}\det P}}\right) \exp\left(-\frac{1}{4}\vec{\sigma}_1^T P^{-1}\vec{\sigma}_1\right) \exp\left(-\frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$$

$$= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\det P P}} \exp\left(-\frac{1}{4}\vec{\sigma}_1^T P^{-1}\vec{\sigma}_1\right) \int_{-\infty}^{\infty} \exp(V\vec{\beta}^T\vec{1}) \times \left(\sqrt{\frac{1}{V^{2(N-1)}}}\right) \exp\left(-\frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$$

$$= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\det P}} \exp\left(-\frac{1}{4}\vec{\sigma}_1^T P^{-1}\vec{\sigma}_1\right) \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$$

$$= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\Pi_a \lambda_a}} \exp\left(-\frac{1}{4}\vec{\sigma}_1^T P^{-1}\vec{\sigma}_1\right) \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$$

where λ_a are the eigenvalues of P

$$\Rightarrow -\log(\mathscr{Z}) = \text{constants} + \frac{1}{2}\sum_a \log(\lambda_a) + \frac{1}{4}\vec{\sigma}_1^T P^{-1}\vec{\sigma}_1 - \log(f(P))$$

where $f(P) = \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$.

$$\Rightarrow \frac{\partial - \log(\mathscr{Z})}{\partial P_{ij}} = \frac{1}{2}\sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4}\sum_{kl} \sigma_{1k} P^{-1}_{ki} P^{-1}_{jl} \sigma_{1l} - \frac{\partial \log f}{\partial P_{ij}}$$

$$\Rightarrow \frac{\partial - \log(\mathscr{Z})}{\partial P_{ij}} = \frac{1}{2}\sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4}\sum_{kl} \sigma_{1k} P^{-1}ki P^{-1}_{jl} \sigma_{1l} - \frac{1}{f}\frac{\partial f}{\partial P_{ij}}$$

$$\frac{\partial f}{\partial P_{ij}} = \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) \left(-\frac{1}{4V}\sum_{kl} \sigma_{2k} P^{-1}_{ki} P^{-1}_{jl} \sigma_{2l}\right) dV$$

$$= \left(-\frac{1}{4}\sum_{kl} \sigma_{2k} P^{-1}_{ki} P^{-1}_{jl} \sigma_{2l}\right) \int_{-\infty}^{\infty} \frac{1}{V^{(N)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV$$

$$\Rightarrow \langle z_i z_j \rangle = \frac{1}{2}\sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4}\sum_{kl} \sigma_{1k} P^{-1}ki P^{-1}_{jl} \sigma_{1l} + \left(\frac{1}{4}\sum_{kl} \sigma_{2k} P^{-1}_{ki} P^{-1}_{jl} \sigma_{2l}\right) \frac{\int_{-\infty}^{\infty} \frac{1}{V^{(N)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV}{\int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp\left(V\vec{\beta}^T\vec{1} - \frac{1}{4V}\vec{\sigma}_2^T P^{-1}\vec{\sigma}_2\right) dV}$$

6.3 Equivalence of dynamics and Hamiltonian

Claim 2 If $H(\vec{v}) = -\vec{v}^T A \vec{v} - \vec{\beta}^T \vec{v}$ and $\frac{d\vec{v}}{dt} = A\vec{v} + \vec{\beta}$ then $\nabla H \cdot \frac{d\vec{v}}{dt} \leq 0$.

Proof.

$$H(\vec{v}) = -\sum_{ij} A_{ij} v_i v_j - \sum_i \beta_i v_i$$

$$\Rightarrow \frac{\partial H}{\partial v_i} = -\sum_{j \neq i} A_{ij} v_j - \sum_{j \neq i} A_{ji} v_j - 2 A_{ii} v_i - \beta_i$$

$$= -\sum_j (A_{ij} + A_{ij}^T) v_j - \beta_i$$

$$\Rightarrow \nabla H = -(A + A^T)\vec{v} - \vec{\beta}$$

$$\Rightarrow \nabla H \cdot \frac{d\vec{v}}{dt} = -\vec{v}^T (A + A^T) A \vec{v} - \vec{v}^T (A + A^T)\vec{\beta} - \vec{\beta}^T A \vec{v} - \vec{\beta}^T \vec{\beta}$$

$$= -\vec{v}^T A^2 \vec{v} - \vec{v}^T A^T A \vec{v} - \vec{v}^T A \vec{\beta} - 2\vec{v}^T A^T \vec{\beta} - \vec{\beta}^T \vec{\beta}$$

If A is negative semi-definite, then A^2 is positive semi-definite and $A^T A$ is positive semi-definite. (Consider that $\vec{v}^T A^T A \vec{v} = (A\vec{v})^T (A\vec{v})$ so that $\vec{v} A^T A \vec{v} = 0$ if $A\vec{v} = 0$ and $\vec{v}^T A^T A \vec{v} < 0$ otherwise.) Therefore, $-\vec{v}^T A^2 \vec{v} \leq 0$ and $-\vec{v}^T A^T A \vec{v} \leq 0$. At equilibrium, $\vec{v}$ will be "close" to $\vec{\beta}$ so that

$$-\vec{v}^T A \vec{\beta} - 2\vec{v}^T A^T \vec{\beta} \sim -\vec{\beta}^T A \vec{\beta}$$

■

6.4 Numerically evaluating dynamics

Claim 3 Let $\Phi(t)$ be the fundamental matrix solution to the homogeneous differential equations $\frac{d\vec{x}}{dt} = A\vec{x}$ so that $\Phi(t) = V\Lambda(t)$ where the columns of V are the eigenvectors of A and

$$\Lambda(t) = \operatorname{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})$$

where $\{\lambda_i\}$ are the eigenvalues of A. (Note that if $\lambda_i = 0$ then $e^{\lambda_i t} = 1$ for all t.) Let $V^\dagger$ be the pseudoinverse of V. Let $\vec{c} = V^\dagger \vec{X}_0$. Let

$$G(t) = \operatorname{diag}(1/\lambda_1 (1 - e^{-\lambda_1 t}), \ldots, 1/\lambda_n (1 - e^{-\lambda_n t})) V^\dagger \vec{\beta}.$$

(If $\lambda_i = 0$ then the corresponding entry of the first part of $G(t)$ will be given by t rather than the form above.) THEN if we let $x(t) = \Phi(t)G(t) + \Phi(t)\vec{c}$, $X(t)$ solves the inhomogeneous equations $\frac{dx}{dt} = A\vec{x} + \vec{\beta}$ with initial conditions $x(0) = \vec{X}_0$.

Proof. It is clear that

$$\frac{d\Phi\vec{c}}{dt} = A\Phi(t)\vec{c} \text{ and } \Phi(0)\vec{c} = \vec{X}_0.$$

Now,

$$\Phi(t)G(t) = V \begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} \frac{1}{\lambda_1}(1 - e^{\lambda_1 t}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda_n}(1 - e^{\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta}$$

$$= V \begin{pmatrix} \frac{1}{\lambda_1}(e^{\lambda_1 t} - 1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda_n}(e^{\lambda_n t} - 1) \end{pmatrix} V^\dagger \vec{\beta}$$

$$\Rightarrow \frac{d\Phi(t)G(t)}{dt} = V \begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix}$$

$$= V \begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix} V^\dagger \vec{\beta} - VV^\dagger \vec{\beta} + \vec{\beta}$$

$$= V \begin{pmatrix} e^{\lambda_1 t} - 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} - 1 \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta}$$

$$= V \begin{pmatrix} \lambda_1 e^{\lambda_1 t} \cdot \frac{1}{\lambda_1}(1 - e^{-\lambda_1 t}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n e^{\lambda_n t} \cdot \frac{1}{\lambda_n}(1 - e^{-\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta}$$

$$= A V \Lambda(t) \begin{pmatrix} \frac{1}{\lambda_1}(1 - e^{-\lambda_1 t}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda_n}(1 - e^{-\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta}$$

$$= A V \Lambda(t) G(t) + \vec{\beta}$$

$$= A \Phi(t) G(t) + \vec{\beta}$$

Therefore $\frac{dx}{dt} = A\Phi(t)G(t) + A\Phi(t)\vec{c} + \vec{\beta} = Ax(t) + \vec{\beta}$ and since $G(0) = \vec{0}$, $x(0) = \vec{X_0}$. ∎

References

[1] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Science*, 105(4):1232–1237, January 2008.

[2] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Oliver Pohl, Edmondo Silvestri, Massimiliano Viale, and Aleksandra Walczak. Social interactions dominate speed control in driving natural flocks toward criticality. *arXiv*, 2013.

[3] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale, and Aleksandra M. Walczak. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Science*, 109(13):4786–4791, March 2012.

[4] Roy J. Glauber. Time dependent statistics of the ising model. *Journal of Mathematical Physics*, 4(2):294–307, 1963.

[5] Petter Holme and Gourab Ghoshal. Dynamics of networking agents competing for high centrality and low d. *Phys. Rev. Lett.*, 96(4), March 2006.

[6] Carl Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 2001.

[7] Martin Nowak. *Evolutionary Dynamics*. Belknap / Harvard, 2006.

[8] George F. Young, Luca Scardovi, Andrea Cavagna, Irene Giardina, and Naomi E. Leonard. Starling flock networks manage uncertainty in consensus at low cost. *PLOS Computational Biology*, 9, 2013.

[9] George Forrest Young, Luca Scardovi, and Naomi Ehrich Leonard. Robustness of noisy consensus dynamics with directed communication. *Proceedings of the American Control Conference*, 2010.