Separating out research-related tidbits from non-research ones.

Contents

1 06/29/19—Collisionless Boltzmann Equation in Galaxies: Landau Damping 1
1.1 Linearized EOM 1

2 02/16/23—Linear Predictive Coding: Autoregressions and Fourier Transforms 2
2.1 Intuitive Understanding 3

3 02/21/23—Chaotic vs Diffusive Behavior 3

1 06/29/19—Collisionless Boltzmann Equation in Galaxies: Landau Damping

Inspired by `https://arxiv.org/pdf/1906.08655.pdf`. The problem is basically formulated as thus: consider a kinetic-theoretic description of a fluid using distribution function $f(t,x,p)$ which obeys collisionless Boltzmann equation $\frac{\mathrm{d}f}{\mathrm{d}t} = 0$ (we use p instead of v to work in Hamiltonian coordinates). Introducing a periodic perturbation to this fluid results in a singular dispersion relation, which can be resolved via the usual Landau prescription (consider a perturbation having grown from zero at $t = -\infty$). The dispersion relation describes *Landau damping* (or growth), in which energy from the fluid is exchanged with the perturber.

1.1 Linearized EOM

The point of the paper is instead to analytically compute the impact of the perturber on the distribution function, to quantify the *scarring* of a galaxy upon encounters with a nearby perturber. The equations of motion coupling the distribution function and gravitational potential are given

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \{*\} f, \mathcal{H} = 0, \tag{1}$$

where $\mathcal{H} = \frac{p^2}{2} + \Phi$ and $\{*\}\ldots$ denotes the Poisson bracket $\{*\} f, \mathcal{H} = \vec{\nabla}_x f \cdot \vec{\nabla}_p \mathcal{H} - \vec{\nabla}_p f \cdot \vec{\nabla}_x \mathcal{H}$.

If we linearize for perturbation quantities f_1, Φ_1 where $\Phi_1(x)$ does not depend on the momenta, we obtain

$$\begin{aligned}
0 &= \frac{\partial f_1}{\partial t} + \{*\} f_1, \mathcal{H}_0 - \vec{\nabla}_p f \cdot \vec{\nabla}_x \mathcal{H}_0, \\
&= \frac{\partial f_1}{\partial t} + \vec{\nabla}_x f_1 \cdot \vec{p} - \vec{\nabla}_p f_1 \cdot \vec{\nabla}_x \Phi_0 - \vec{\nabla}_p f_0 \cdot \vec{\nabla}_x \Phi_1.
\end{aligned}$$

Oops welp I guess I never solved this.

2 02/16/23—Linear Predictive Coding: Autoregressions and Fourier Transforms

This was a simple enough inquiry initially: given a partial time series that contains sinusoids, how do we extract the frequency? We know one way to do this using the FFT, but there are advantages to other techniques. Courtesy of Jeremy Goodman's pointers.

The trick has to do with autoregressions. Suppose we are looking to extract l frequencies of form $e^{i\omega_m t}$, so that

$$y_n = \sum_m^l C_m e^{i\omega_m n\Delta t}. \tag{2}$$

Thus, if we have at least $2l$ points or so, we should be able to fit for the $2l$ DOF C_m and ω_m. There can be a noise term above if need be, in which case more points will smooth out the noise.

What is the trick? Well, we compute the l-th order autoregression. In other words, for each sequence of l points, we can write down the expression satisfying:

$$y_n = \sum_{m=1}^{l} a_m y_{n-m}. \tag{3}$$

With l-many such sequences, we have enough equations to solve for the l-many unknowns a_m. These can be written in the matrix form:

$$\begin{bmatrix} y_n \\ y_{n+1} \\ \vdots \\ y_{n+l} \end{bmatrix} = \begin{bmatrix} y_{n-1} & y_{n-2} & \cdots & y_{n-l} \\ y_n & y_{n-1} & \cdots & y_{n-l+1} \\ \vdots & \vdots & \ldots & \vdots \\ y_{n+l-1} & y_{n+l-2} & \cdots & y_{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{bmatrix}. \tag{4}$$

These $\{a_l\}$ form the AR(l) autoregressive model for y_n.

This is great, but how do we get the frequencies, or also maybe the growth rates? Now, we rewrite the above equation as

$$0 = \begin{bmatrix} y_n & y_{n-1} & \cdots & y_{n-l} \\ y_{n+1} & y_n & \cdots & y_{n-l+1} \\ \vdots & \vdots & \ldots & \vdots \\ y_{n+l} & y_{n+l-1} & \cdots & y_{n-1} \end{bmatrix} \begin{bmatrix} -1 \\ a_1 \\ \vdots \\ a_l \end{bmatrix} \equiv \mathbf{B} \cdot \vec{a}. \tag{5}$$

Now, what if the y_n look like λ^n for some complex λ? Then the λ must satisfy

$$1 - a_1\lambda - a_2\lambda^2 - \cdots - a_l\lambda^l = 0. \tag{6}$$

This is the characteristic equation for this AR(l) model. If we solve for the roots of this equation, we get the possible values of λ that satisfy the model. In other words, if the $y_n = \lambda^n$ indeed, then $\mathbf{B} \cdot \vec{a} = 0$ as requested above. Then, if the data are oscillatory, then $\lambda = e^{i\omega_m}$ as requested above.

2.1 Intuitive Understanding

There's something slightly unintuitive here: we began by seeking the frequencies in the y_n, but we eventually obtained this by solving an equation that has *nothing* to do with the y_n, the characteristic equation for the AR(l) model! Why does this make sense?

Well, it should be noted that a particular AR(l) model does not uniquely specify the y_n; this would be impossible, since there are only l DOF in the model and $2l+1$ in Eq. (4). Indeed, this suggests that the amplitudes of the modes C_m, as well as the initial normalization of the autoregressive chain (y_{n-l}) are free parameters. As such, we can imagine that the AR(l) model permits a family of solutions, any with the correct frequencies. In other words, we could also imagine writing:

$$\mathbf{B} \cdot |a\rangle = \sum_{m=1}^{l} C_m |b_m\rangle \langle b_m| |a\rangle . \tag{7}$$

Another way to think about the characteristic equation is as exactly a characteristic equation of a matrix. If we consider the matrix that maps the vector

$$\begin{bmatrix} y_N \\ y_{N-1} \\ \vdots \\ y_{N-l} \end{bmatrix} = \mathbf{M} \begin{bmatrix} y_{N-1} \\ y_{N-2} \\ \vdots \\ y_{N-l-1} \end{bmatrix}, \tag{8}$$

then it's clear that $\mathbf{M}$ has the form

$$\mathbf{M} = \begin{bmatrix} a_1 & a_2 & \dots & a_{l-1} & a_l \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \tag{9}$$

It's then clear $\mathbf{M}$ apparently has exactly the characteristic equation that we prescribe above. This makes sense: the matrix $\mathbf{M}$ tells us whether a vector-valued sequence of y values is growing, decaying, or oscillating.

As such, the final conclusion of this tidbit is this: the autoregression is another way of expressing a Markov chain that allows us to advance the time series. Then, note that any sequence with $y_n = z^n$ where z is complex (allowing periodic or exponential sequences) has z as one of the eigenvalue of its Markov chain matrix, or z is a root of its characteristic polynomial. Turning this on its head, if we compute the autoregression for a sequence and find a root w of its characteristic polynomial, this implies that the sequence has a geometric component with factor w. Applying this to sequence with a periodic component with frequency ω, we see that $e^{i\omega\Delta t}$ must be a root of the characteristic polynomial of its autoregression.

3 02/21/23—Chaotic vs Diffusive Behavior

This is a short section. Dong (and others) talk about "chaotic tides", where the mode amplitude grows stochastically because the forcing occurs with random amplitude. However, this is not chaos, but should be properly termed "diffusive tides."

How can we argue for this? Well, the defining characteristic of chaos is a positive Lyapunov exponent, i.e. an exponential growth rate of the separation between two trajectories with nearly-idential initial conditions.

$$\delta y(t) = y(t; y_0) - y(t; y_0 - \epsilon), \tag{10}$$

$$\sim \mathcal{O}(\epsilon e^{\lambda t}). \tag{11}$$

What is the growth rate for a random walk, or diffusive growth?

Let's adopt the simplest model for now, a discrete random walk with step size ± 1. Perhaps, for the sake of consistency, we can imagine that the step is determined based on the current value of x, e.g. whether the rounded value of $10^9 x$ is even or odd. Then, consider two initially adjacent x. It is obvious that

$$|\delta x(t)| \leq 2t. \tag{12}$$

So we already see that diffusive behavior is not chaotic.

But now, we have a fun little math problem. Consider two random walks starting at $x = 0$ with step size ± 1. What is the mean and variance of δx? Well, using the usual CLT guidelines, the linearity of expectation gives $\mathrm{E}(X_2 - X_1) = 0$ while linearity of variance gives $\mathrm{Var}(X_2 - X_1) = 2\,\mathrm{Var}(X_1) = 2t$. Thus, the separation between two walkers grows stochastically and $\sim \sqrt{t}$. This is not chaos, where the separation grows deterministically and $\sim \exp(\lambda t)$.