Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

SEMIPARAMETRIC INFERENCE FOR SYNCHRONIZATION OF POPULATION CYCLES

P.E. Greenwood
Dept. of Mathematics, Arizona State University, Tempe, Arizona
and Dept. of Mathematics, University of British Columbia, Canada
D.T. Haydon
Centre for Tropical Veterinary Medicine
Easter Bush, Roslin, Midlothian, U.K.

1 Introduction

We consider a dynamic random field. On each of a discrete array of sites is located a hyperbolic dynamical system perturbed by noise. We assume that the dynamics are identical, with a unimodal limit cycle and that the perturbations are independent, centered, and with the same distribution. In addition the individual processes are coupled with one another in a homogeneous pattern. The coupling may be global, in which case we are thinking of a mean-field type of system. Or the coupling may be local, the coupling strength between each site and its neighbors attenuating with distance.

The application we have in mind is to cycling populations of animals, where the log of each local population increases roughly linearly to an apparent critical point from which it falls precipitously to a minimum. The data is discrete in space and time, being based on periodic reports from catchment or reporting regions. In a previous study [1] we focussed on data from Canadian lynx populations. Lynx population cycles are known to follow those of snowshoe hare. Previous analyses of this data have been concerned with inferring the length and regularity of the evident population cycles. We consider in [1] the very different challenge of estimating a parameter identified as strength of coupling among populations. That paper is primarily addressed to data analysis and interpretation.

The emphasis in this report is on the steps involved in arriving at a suitable model and estimator. We explore some of the difficulties posed by this rather unusual problem. In forthcoming studies we, with colleagues, will apply the method described here to cycling population data from Canadian muskrat and mink, and from the greysided vole of Hokkaido.

2 Development of a Coupled Evolving Phase Field Model

To begin, let us introduce a general definition of synchronization applicable to random fields. Let $X = \{X_{it}, i \in I, t = 0, 1, 2, \ldots\}$ denote the values at site i and time t of a dynamic random field. Let $S = \{S(\cdot)_t, t \geq 0\}$ denote a real-valued functional or collection of functionals of $\{X_{is}, i \in I, s \leq t\}$. Note that we allow S evaluated at time t to depend on data up to and including time t. Define S-synchronization to mean $(SX)_t \equiv 0$, almost surely, for t in some time-set T. This gives us a flexible definition of exact synchronization. But a random field will not be exactly synchronized. In order to formulate a statistically useful definition we need to allow for and to measure departures from synchronization. For example, we probably want the mean and variance of S applied to random data to be small. We are led to the following:
Definition. A dynamic random field X departs from S-synchronization by no more than ϵ over the time interval T if $E(SX)_t^2 \leq \epsilon$ for all $t \in T$.

One might use other norms, e.g. L^1 or Kullbach-Leibler instead of L^2. However this definition has a familiar form and is computationally convenient.

At each site i in an array of sites I is located a "skeleton" process which, in the absence of any noise or coupling can be written

$$X_{i,t+1} = f(X_{it}), \quad i \in I, t = 0, 1, 2, \ldots. \tag{2.1}$$

In this treatment we write the dynamics as of first order in discrete time. We assume there is a unimodal limit cycle $\ell(t), t = 0, \ldots, p$, $\ell(0) = \ell(p)$, where, since f is the same for each i, the form of the cycle, ℓ, and its period, p, are the same for each site i. For definiteness let $\ell(0)$ be the minimum point of the unimodal cycle. If we picture the deterministic field (1) running in equilibrium, we see the periodic cycle $\ell(\cdot)$ being executed at each site, the only difference between sites being the phase, which will differ if the initial points $X_{i,0}$ were not identical. Now suppose we watch, instead, the stochastic field,

$$X_{i,t+1} = f(X_{it}) + \epsilon_{it}, \quad i \in I, t = 0, 1, 2, \ldots \tag{2.2}$$

where ϵ_{it} represents a small, identically distributed, centered noise. A theorem of deterministic dynamics called the "shadowing lemma" implies that the stochastic paths of (2) shadow the paths of (1), where "shadow" means that they remain in a distributional neighborhood of the deterministic path up to a time shift. The differences among the paths at the various sites, then, in addition to the small width of this neighborhood, are the phases, which change in a random way. The neighborhood width depends on the noise variance. This picture motivates a model for coupling based on the randomly evolving phases of the components.

We now move our attention from the evolving random field of population levels, (2), to a corresponding evolving random field of phases. Going back to the deterministic model (1) we can unambiguously define the phase ϕ_{it} at site i, time t, to be the time fraction of the current cycle, $\ell(0), \ldots, \ell(p)$, which has been accomplished at time t. Then each phase ϕ_{it} is in the interval $[0, 1)$. The arithmetic for phases is mod1. We think of the set of points $\{\phi_{it}, i \in I\}$ as a set on the circle of circumference 1. As time advances the points progress around the circle. Since in fact we wish to consider the phase field associated with the stochastic field (2), where the "limit cycle" is a stochastic perturbation of the deterministic one, we may not see a unique minimum for each cycle, and the definition of ϕ_{it} may be ambiguous. We will assume that the noise ϵ_{it} is small enough so that the ambiguity can be resolved by a device described in the next section. For the moment let us ignore this problem and assume that in the stochastic model the phase ϕ_{it} is defined as the fraction of the current orbit which has been traversed at time t by each path X_i of the stochastic field (2). We describe the structure of the phase field $\phi_{it}, i \in I, t = 0, 1, 2, \ldots$, by writing

$$\phi_{i,t+1} = \phi_{it} + g_{it} + \epsilon_{it} \quad \bmod 1, \tag{2.3}$$

where g_{it} is the fraction of the current orbit traversed at site i at time t.

Now we introduce a hypothetical coupling force into the phase field which will shift the phase at each site i and time t in the direction of the "mean phase". For this purpose we need to devise an appropriate definition of mean for a set of random points on a circle.

As mentioned above we identify the phase values with points on a circle of circumference 1. Between points x, y on this circle, let $\Delta(x, y)$ denote the signed smallest arc measured counter-clockwise between them. Then $\Delta(x, y)$ is positive or negative according to whether x leads or lags y, and $|\Delta(x, y)| \leq 0.5$.

For each set of phases ϕ_{it} with t fixed, let $\overline{\phi}_t$ be a solution of

$$\sum_{i \in I} \Delta(\phi_{it}, \overline{\phi}_t) = 0. \tag{2.4}$$

Then $\overline{\phi}_t$ is almost surely uniquely defined and can be interpreted as the mean phase at time t on the circle, or equivalently on $[0, 1)$ mod 1. To model phase coupling among sites i in the array I, we insert a coupling term in expression (3) which moves the phases at time $t + 1$ in the direction of the mean phase at time t, and write

$$\phi_{i,t+1} = \phi_{it} + g_{it} - c\Delta(\phi_{it}, \overline{\phi}_t) + \epsilon_{it} \bmod 1, \quad i \in I, t = 0, 1, \ldots. \tag{2.5}$$

The coupling strength c can take values in $[0, 1)$. Expression (5) says that in each time step, from t to $t + 1$, three changes occur in the phase at each site:

(a) it advances according to phase dynamic g_{it} inherited from the limit cycle of the "skeleton" process (1),

(b) it is shifted in the direction of the mean of the phases as defined by (4) by a proportion c of its distance from this mean, and

(c) it is perturbed by a centered random effect.

About the ϵ_{it} we assume only that for each t they are conditionally independent given the past of the process $\{\phi_{it}\}$, with the same conditional distribution for each i and t.

In order to simplify our computations we would like to omit the mod1 in relation (5). This is possible if it does not happen, or happens only rarely, that the combined effect of (a), (b), and (c) at a particular time-step t is large enough that the phase increment at a single step exceeds 1. In the model we make this assumption. If it is occasionally violated in the data we make an appropriate adjustment. Assuming, then, that the noise is sufficiently small relative to the coupling, (5) can be written as

$$\begin{aligned} \Delta(\phi_{i,t+1}, \overline{\phi}_{t+1}) &= (1-c)\Delta(\phi_{it}, \overline{\phi}_t) + \epsilon_{it} + (g_{it} - \overline{\phi}_{t+1} + \overline{\phi}_t), \\ & \quad i \in I, t = 0, 1, \ldots. \end{aligned} \tag{2.6}$$

The last grouped term in (6) has mean 0 as can be seen from considerations of symmetry. In addition, for each t the various values in i are conditionally independent given the past of the phase field. We have assumed that the noise increments ϵ_{it} are small and centered and conditionally independent given the past of the phase field. Without altering this assumption we can regard the last grouped term in (6) as part of the noise and write (6) as

$$\Delta(\phi_{i,t+1}, \overline{\phi}_{t+1}) = (1-c)\Delta(\phi_{it}, \overline{\phi}_t) + \eta_{it}, \quad i \in I, t = 0, 1, \ldots. \tag{2.7}$$

Relation (7) describes a stochastic process of AR1-type. It differs from the usual AR1 model in that the process is a vector-valued function of a distinct vector process. The underlying process has vectors of phases as its values, whereas the process defined by (7) has values which are vectors of differences of these phases from the evolving central phase. Nevertheless the least squares type estimator is consistent for $(1 - c)$ and that is what we used in [1]. One needs to know that there is a stationary law for the process defined by (7). This can be shown by an argument along the lines of arguments in Meyn and Tweedie's book [2]. The fact that the state space is bounded simplifies the situation.

If we wish to model local rather than global coupling, then the central phase as seen from site i should be computed from the phases at sites near i. We then define the local centre of phases at i, to be the almost surely unique phase $\overline{\phi}_{it}$ which is the solution of

$$\sum_j a_{ij}\Delta(\phi_{jt},\overline{\phi}_{it}) = 0, \quad i \in I, t = 0, 1, \ldots, \tag{2.8}$$

where the numbers a_{ij} are weights reflecting the fraction of the total coupling at site i to be attributed to interaction between sites i and j, $\sum_j a_{ij} = 1$. We arrive at a version of (7) with $\overline{\phi}_{it}$ replacing $\overline{\phi}_t$ in each equation. If edge effects are disregarded, and if a_{ij} depends only on $i - j$, the system remains spatially homogeneous in law. In any case, a stationary distribution exists for the process and the least-squares-type estimator is consistent for $1 - c$.

Let us return to the notion of departure from S-synchronization by no more that ϵ, as defined early in this section. Perfect synchronization corresponds to departure zero, and higher levels of departure correspond to less synchronization. The synchronization measure we choose here is the mean-square deviation of the phase field from its mean $\overline{\phi}_t$, or $\overline{\phi}_{it}$, as defined by (4) or (8),

$$(SX)_t^2 = \frac{1}{|I|}\sum_{i\in I}\Delta(\phi_{it},\overline{\phi}_t)^2. \tag{2.9}$$

The departure from synchronization, then, is defined as

$$E(SX)_t^2 = \frac{1}{|I|}\sum_{i\in I}E\Delta(\phi_{it},\overline{\phi}_t)^2. \tag{2.10}$$

3 Estimation

In an initial application of the model presented here [1], we study synchronization in the well-known Canadian lynx data sets. Here we summarize a few of the results and refer to [1] for the full treatment and references.

The first data set was compiled by Elton and Nicholson, in 1942, [4] spans the years 1821 to 1891 and is organized over six trading regions of the Hudson's Bay Company. The second data set, compiled by Statistics Canada, [5] is organized over eight Canadian provinces and territories and spans the years 1919–1990. We used estimators of $(1 - c)$ and σ_η^2 where η is the noise increment in (7), which are analogous to well-known linear estimators for parameters of AR1 time series.

The data analysis involves two difficulties peculiar to the phase-coupled model (7). As mentioned earlier it was not possible to identify uniquely the time points where minima of cycles occur in the data. Our treatment in [1] involves identifying the sets of possible minima and repeating the estimation procedure many times using random choices from these sets. We are exploring other possible ways to handle this problem in other data sets. A second difficulty involves discontinuities in the function $\Delta(\phi_{it}, \overline{\phi}_t)$ at ±0.5, and a very few instances where $\Delta(\overline{\phi}_{t+1}, \overline{\phi}_t) < 0$. Our handling of these is detailed in [1]. We mention these problems here to indicate the novel difficulties posed by data analysis in this context.

The estimators we use for $\theta = 1 - c$ and for σ_η^2 are

$$\begin{aligned} \hat{\theta} &= \frac{\sum_{i=1}^{n}\sum_{t=1}^{N-1} \Delta(\phi_{it}, \overline{\phi}_t)\Delta(\phi_{i,i+1}, \overline{\phi}_{t+1})}{\sum_{i=1}^{n}\sum_{t=1}^{N-1} \Delta(\phi_{it}, \overline{\phi}_t)^2}, \\ \hat{\sigma}_\eta^2 &= \left(\sum_{i=1}^{n}\sum_{t=1}^{N-1} \left[\Delta(\phi_{i,t+1}, \overline{\phi}_{t+1}) - \overline{\Delta(\phi_{i,t+1}, \overline{\phi}_{t+1})}\right]^2 \right. \\ &\quad \left. -\theta^2 \sum_{i=1}^{n}\sum_{t=1}^{N-1} \left[\Delta(\phi_{it}, \overline{\phi}_t) - \overline{\Delta(\phi_{it}, \overline{\phi}_t)}\right]^2\right) / n(N-1) - 2. \end{aligned} \tag{3.1}$$

We can distinguish by looking at the residuals from these estimates the times at which synchrony is maintained by intermittent synchronizing events from the time periods of constant phase-coupling dynamics as expressed in expression (7). At the times of synchronizing events we see outliers in the residuals, whereas corresponding to periods of consistent coupling, data is clustered around the regression line. Times, thus detected, of synchronizing events correspond to times of decrease of the estimator (9) of asynchrony.

The coupling estimates we obtained were $\tilde{c} \sim 0.054$ for the earlier data set and $\tilde{c} \sim 0.011$ for the later one. With local weighting the estimate doubled to $\tilde{c} \sim 0.096$ for the earlier data set and remained about the same at $\tilde{c} \sim 0.005$ for the later data set. The numbers for the earlier data set are significant according to tests using the estimated variances, and are significantly greater than estimates produced from simulated cycles with randomized phases from a model fitted to this data. Thus, although the estimates $\tilde{c}$ from the lynx data seem small, we believe that they measure the phenomenon of coupling as intended.

An alternative statistical approach to the study of synchronization of population cycles uses the correlation structure of the logged abundance data as in e.g. Jolliffe [3]. With this method, amplitude variability may confound the detection of synchronization. Further, since phase coupling is not modelled, correlational methods measure phase coupling only very indirectly.

Acknowledgements

This work was supported by NSERC Canada and the Crisis Points Group of the Peter Wall Institute of Advanced Studies at U.B.C. The second author acknowledges support from the Wellcome Trust.

References

[1] Haydon, D.T. and Greenwood, P.E. (2000). Spatial coupling in cyclic population dynamics: models and data. *Theoretical Population Biology*, **58**, 239-254.

[2] Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability.* Springer, Heidelberg.

[3] Jolliffe, I.T. (1986). *Principal Components Analysis.* Springer, New York.

[4] Elton, C. and Nicholson, M. (1942). The ten-year cycle in numbers of lynx in Canada. *J. Anim. Ecol.*, **11**, 215-244.

[5] Statistics Canada. 1921-1990. Catalogue **23-207**, Fur Production, Ottawa, Canada.