Analyzing Stochastic Diffusion Processes

Introduction

- Many interesting ecological diffusions
- Emerging diseases - avian flu, H1N1 flu
- Exotic organisms - invasive plants, gypsy moths
- Size and age distributions
- Transformation of landscape, deforestation, land use classifications, urban growth

The objective

- Our objective: forecast likely spread in space and time with associated uncertainty
- Nonlinear, nonhomogeneous in space and time
- Explanatory covariates
- Start with deterministic integro-differential equations or with partial differential equations
- How to add uncertainty?

cont.

- Theoretical models, “varying” around them
- Too much simplification required to obtain analytical solutions
- Discretization to fit models
- Do we care of the deterministic equation? Should we just work with the discrete time version we want? Dynamic spatial models?

Hierarchical modeling

- The hierarchical paradigm:

$$[data|process, parameters][process|parameters][parameters]$$

- A paradigm shift - designed experiments to observational studies; controlled experiments to integrated (big picture) investigation
- Prior information from: empirical studies, mechanistic knowledge, ecological theory, etc.
- Multiple information sources
- Conditional uncertainty in components (a natural way to specify models)
- Different resolutions in space and time
- Structured dependence in space and time
- Complex dependence structure through latent variables

Computation

- Fit within a Bayesian framework; enables full inference, exact inference
- Model fitting, associated computation is challenging
- High dimension, sparsity, dimension reduction
- MCMC/Gibbs sampler model fitting
- Model validation? model comparison?

Dynamics

- Continuous space, discrete time, i.e., $w_t(s)$
- Without loss of generality $\mathrm{t} = (1, 2, ..., T)$
- Envision $w_t(s)$ as a *dynamical* process
- In fact, simplify to first order Markov, i.e., for locations $s_1, s_2, ..., s_n$, let $\mathbf{w}_t = (w_t(s_1), w_t(s_2), ..., w_t(s_n))^T$. Then

$$[\mathbf{w}_t | \mathbf{w}_0, \mathbf{w}_1, ... \mathbf{w}_{t-1}] = [\mathbf{w}_t | \mathbf{w}_{t-1}]$$

- For example, $\mathbf{w}_t = H\mathbf{w}_{t-1} + \boldsymbol{\eta}_t$ where $\eta_t(s)$ incorporates spatial structure
- A vector AR(1) model and H is called the *propagator* matrix
- ??Specifying H??

Specifying H

- $H = I$ - not stationary (explosive), no interaction across space and time, not realistic for most dynamic processes of interest
- $H =$ Diag(h) where Diag(h) has diagonal elements $0 < h_i < 1$ - Still no interactions
- Integro-difference equation (IDE) dynamics:

$$w_t(s) = \int h(s, r; \phi) w_{t-1}(r) dr + \eta_t(s)$$

- h is a "redistribution kernel" that determines the rate of diffusion and the advection

cont

- If require $w > 0$, work with

$$\log w_t(s) = \log\left(\int h(s, r; \phi) w_{t-1}(r) dr\right) + \eta_t(s)$$

- Alternatively,

$$v_t(s) = \int h(s, r; \phi) v_{t-1}(r) dr$$

and

$$\log w_t(s) = \log v_t(s) + \eta_t(s)$$

- Discretization to obtain H
- Forms for $h(s, r; \phi)$; $h(s, r; \phi(r))$?, $h_t(s, r; \phi)$?

cont

- Recall linear PDE, $\frac{dw(s,t)}{dt} = h(s)w(s,t)$
- Finite differencing yields $w(s,t+\Delta t) - w(s,t) = h(s)w(s,t)\Delta t$, i.e., $w(s,t+1) \approx \tilde{h}(s)w(s,t)$. Same limitations as above.
- Need more general PDE's
- PDE can motivate IDE, can clarify H
- "forward" vs. "backward" perspective
- IDE's can be specified directly without using PDE's, e.g., $h(s,r)$ can be a sum of a survival/spread term $+$ a birth/replenishment term

Diffusion PDE's

- Diffusion in one dimension - Fick's Law: diffusive flux from *high* concentration to *low* is $-\delta\frac{\partial w(x,t)}{\partial x}$ with δ, the diffusion coefficient. Location varying diffusion $\delta(x)$
- And, diffusion equation is $\partial w/\partial t = -\partial\text{flux}/\partial x$, i.e., $\frac{\partial w(x,t)}{\partial t} = \frac{\partial}{\partial x}(\delta(x)\frac{\partial w(x,t)}{\partial x})$
- That is, the 1-dim diffusion equation is

$$\frac{\partial w(x,t)}{\partial t} = \delta'(x)\frac{\partial w(x,t)}{\partial x} + \delta(x)\frac{\partial^2 w(x,t)}{\partial x^2}$$

- In 2-dim, diffusive flux is $-\delta(x,y)\nabla w(x,y,t)$ ($\nabla w(x,y,t)$ is the concentration gradient at time t)
- The resulting diffusion PDE is

$$\frac{\partial w(x,y,t)}{\partial t} = \frac{\partial}{\partial x}(\delta(x,y)\frac{\partial w(x,y,t)}{\partial x}) + \frac{\partial}{\partial y}(\delta(x,y)\frac{\partial w(x,y,t)}{\partial y})$$

Discretizing the diffusion equation

- Complete the differentiation of the diffusion equation
- Yields second order partial derivatives, $\frac{\partial^2 w}{\partial x^2}$ and $\frac{\partial^2 w}{\partial y^2}$
- Introduce Δt, Δx, Δy
- Replace ∂'s with finite differences (first forward and second order centered) - careful detail, ugly expression!
- After the smoke clears, we obtain $\mathbf{w}_{t+\Delta t} = H\mathbf{w}_t$
- Again, add $\boldsymbol{\eta}_t$
- We are back to our earlier redistribution form

Add growth rate

- Previous dynamics simply redistribute existing population spatially over time
- In many situations, there is also growth of the population
- Population growth can be captured by a logistic differential equation

$$\frac{\partial w(s,t)}{\partial t} = rw(s,t)(1 - w(s,t)/K)$$

- r is the growth rate, K is the carrying capacity
- Add growth to the diffusion PDE for $\frac{\partial w(s,t)}{\partial t}$
- $r(s)$?, $K(s)$?

Eurasian collared dove data

- An example from Wikle et al., using data from the Breeding Bird Survey (BBS)
- Escaped to U.S. from Bahamas, introduced in Florida, expanding dramatically across North America
- $4000+$ routes in the survey (some sampled more than once per year, others not sampled in a given year), length of route is ≈ 40 kms, 50 stops per route, count birds by sight for 3 minutes item 18 years: 1986-2003
- Route is a “point”, response at a point is a count
- Aggregate to grid boxes
- Z_{it} is count in box i in year t, n_{it} is number of visits to cell i in year t.
- λ_{it} is *intensity* for box i in year t

Modeling specifics

- $Z_{it} \sim \text{Po}(n_{it}\lambda_{it})$
- $\log \lambda_{it} = w_{it} + \epsilon_{it}$
- ϵ_{it} are i.i.d. (pure error or micro-scale variation)
- The focus is on the $\mathbf{w}_t$. They tell the diffusion story, i.e., $\mathbf{w}_t = H(\boldsymbol{\delta})\mathbf{w}_{t-1} + \boldsymbol{\eta}_t$
- Model for $\boldsymbol{\eta}_t$?
- $\mathbf{w}_0 \sim N(0, 10I)$
- δ is the vector of local diffusion coefficients, one for each grid cell
- A dimension reduction for δ; many possibilities here - basis functions, EOF's, predictive processes

Enriching the modeling

- Again, we focus on $w(s,t)$
- $w(s,t)$ can arise as a mean model for a geostatistical model or in a space-time GLM (as in Wikle) or as a *cumulative* intensity $\Lambda(s,t)$ for a space-time point pattern (which drives the cumulative diffusion)
- A general diffusion PDE (nonstochastic) looks like $\frac{\partial w(s,t)}{\partial t} = a(w(s,t), z(s,t), \theta)$ where $z(s,t)$ are other potential variables ($z(s,t) = t$ for example)
- $\theta(s)$?, $\theta(s,t)$?
- How to make the PDE stochastic?
- For the remainder, we use the logistic DE, i.e., a diffusion driven by a logistic growth model
- Discretization as proposed above

First a DE

- Ignoring location s for the moment, we have:

  $$dw(t) = a(w(t), t, \theta)dt \text{ with } w(0) = w_0$$

  Simplest way to add stochasticity is to make θ random.
- Instead:

  $$dw(t) = a(w(t), t, \theta)dt + b(w(t), t, \theta)dZ(t)$$

  where $Z(t)$ is Brownian motion over R^1 with a and b the "drift" and "volatility" respectively. Now a *stochastic* differential equation (SDE)
- θ would still be random

- Next: $dw(t) = a(w(t), t, \theta(t))dt$ where (with $Z(t)$ is variance 1 Brownian motion)

$$d\theta(t) = g(\theta(t), t, \beta)dt + h(\theta(t), \sigma)dZ(t)$$

- This includes the previous example
- For the logistic equation:

$$dw\,(t) = \theta(t)w\,(t)\left[1 - \frac{w\,(t)}{K}\right]dt$$

- If $\theta(t) = \mu + \zeta(t)$ with $d\zeta\,(t) = -a\zeta\,(t)\,dt + \sigma_\zeta dZ\,(t)$ equivalently

$$d\theta\,(t) = -\alpha\,(\mu - \theta\,(t))\,dt + \sigma_\zeta dZ\,(t)\,,$$

a self-reverting Ornstein-Uhlenbeck (OU) process and $\theta(t)$ is a stationary GP with $cov(\theta(t), \theta(t')) = (\sigma^2/\alpha)exp(-\alpha|t - t'|)$.

Add space

- Now, we add space. First,

  $$dw(s,t) = a(w(s,t),t,\theta(s))dt \text{ with } w(s,0)) = w_0(s),$$

  a PDE. Randomness through $\theta(s)$, a process realization, so $\theta(s)$ provides the spatial dependence. Hence, a stochastic process of differential equations.
- Next,

  $$dw(s,t) = a(w(s,t),t,\theta(s))dt + b(w(s,t),t,\theta(s))dZ(s,t)$$

- Modeling $Z(s,t)$? For a fixed finite set of spatial locations assume independent Brownian motion at each location.
- Or a discrete space approximation to spatial Brownian motion employing a Gaussian process (GP)on R^2

cont.

- Next,

$$dw(s,t) = a(w(s,t),t,\theta(s,t))dt$$

  where say

$$d\theta(s,t) = \gamma(\theta(s,t) - \theta(s))dt + bdZ(s,t)$$

- Again, $\theta(s)$ is process realization
- Now, $\theta(s,t)$ given through an infinite dimensional SDE
- This version produces a covariance function that is separable in space and time

Important points

- A differential equation in time at every spatial location, i.e., parameters indexed by location
- The parameters vary spatially as realizations of a spatial process
- Instead, the differential equation is a stochastic differential equation (SDE), e.g., a spatial Ornstein-Uhlenbeck (O-U) process
- For the logistic PDE, the rate parameter in the differential equation is assumed to change over time. It can be modeled as a realization of a spatio-temporal process
- It can be modeled using a SDE, yielding an SDE embedded within the differential equation

- Spatio-temporal modeling settings for the above?
- The usual "geostatistics" setting with observations at locations and times:

$$Y(s,t) = \Lambda(s,t) + \epsilon(s,t)$$

  with $\Lambda(s,t)$ modeled through a differential equation as above, i.e., process model is a stochastic PDE.
- Space-time point pattern setting with data of the form $(s_i, t_i), i = 1, 2, ...n$, i.e., random locations at random times. Now, we model the space-time intensity associated with the point pattern, again, say $\Lambda(s,t)$, through a differential equation as above resulting in a space-time Cox process.

A Geostatistical Example

- The logistic PDE in space and time:

$$\frac{\partial \Lambda(t,s)}{\partial t} = r(t,s)\Lambda(t,s)\left[1 - \frac{\Lambda(t,s)}{K(s)}\right]$$

- Time discretized to intervals Δt, indexed as $t_j, j = 0, 1, 2, ...J$. At location s_i, data $Y(t_j, s_i)$
- Dynamic model: $Y(t_j, s_i) = \Lambda(t_j, s_i) + \varepsilon(t_j, s_i)$
- Using Euler's approximation yields difference equation:
  $\Delta\Lambda(t_j, s) = r(t_{j-1}, s)\Lambda(t_{j-1}, s)\left[1 - \frac{\Lambda(t_{j-1}, s)}{K(s)}\right]\Delta t,$
  $\Lambda(t_j, s) \approx \Lambda(0, s) + \sum_{l=1}^{j} \Delta\Lambda(t_l, s)$
- No spatial flux, no Δx, no Δy, $\Lambda_t = Diag(h)\Lambda_{t-1}$

cont.

- Can not add scaled Brownian motion to the logistic PDE.
- Instead a time-varying growth rate at each location
- An O-U process for $r(t,s)$:

$$\frac{\partial r\left(t,s\right)}{\partial t}=\alpha_r\left(\mu_r\left(s\right)-r(t,s)\right)+\frac{\partial B\left(t,s\right)}{\partial t}$$

- Model the initial $\Lambda\left(0,s\right)$ and $K\left(s\right)$ as log-Gaussian spatial processes with regression forms for the means
- Below we assume $K\left(s\right)$ known and set to 1, i.e., 100% is the capacity.
- Similar modeling for $\mu_r(s)$

A Simulation Example

- 10×10 study region
- 44 locations over 30 time periods
- 4 sites for holdout
- Matérn covariance function used for $\Lambda_0(s)$, for $r(s)$, ($\nu = 3/2$ used in the example)
- E.g., the space time covariance function for $r(t, s)$ becomes $\varrho(t_{j_1} - t_{j_2}, s_{i_1} - s_{i2}) =$ $\sigma_r^2 \exp\left(-\alpha_r \left|t_{j_1} - t_{j_2}\right|\right)\left(\phi_r \left|s_{i_1} - s_{i_2}\right|\right)^\nu \kappa_\nu\left(\phi_r \left|s_{i_1} - s_{i_2}\right|\right)$
- weak priors on σ^2's, weak log normal prior on α_r, discrete prior on ϕ's

Space-time point patterns

- Spatio-temporal Cox process models using SDEŠs
- Motivation: Urban development using spatio-temporal point processes.
- Differential equation models for cumulative intensity
- We have 21 years of urban development data for Irving, TX but we just show a simulation example with five years

Spatio-temporal Cox Process

In a study region D during a period of $[0,T]$, N_T events:

Point pattern: $X_T = \left\{ x_{1,t_1}, \ldots, x_{N_T, t_{N_T}} \right\}$

where $x_{i,t_i} = (x_i^1, x_i^2, t_i)$

X_T is a Poisson process with inhomogeneous intensity

$$\Omega(t,s), s \in D, t \in [0,T]$$

Specifying the intensity?

$$\Omega(t,s) = f\left(t, \theta_l(t,s); l = 1, \ldots, p\right)$$

$\theta_l(t,s), s \in D, l = 1, \ldots, p$ are processes for parameters of interest.

The cumulative intensity

Discretize the spatio-temporal Cox process in time:

Spatial point pattern: $X_{[t_1,t_2]}$ during $t \in [t_1, t_2)$

$$x_i = (x_i^1, x_i^2), x_i \in X_{[t_1,t_2]}$$

The cumulative intensity for $X_{[t_1,t_2]}$ is

$$\int_{t_1}^{t_2} \Omega\,(t,s)\,dt = \int_{t_1}^{t_2} f\,(t, \theta_l(t,s); l = 1, \ldots, p)\,dt$$

We consider models for the cumulative intensity

$$\Lambda\,(t,s) = \int_0^t \Omega\,(\tau, s)\,d\tau$$

Comments

- So house locations and times over $(0, T] \times D$
- Need a Δt and an area A in order to observe a point pattern
- If $\Omega(t, s) \geq 0$ then $\Lambda(t, s)$ increases in t; we do not allow house removal
- Work with cumulative intensity $\Lambda(t, s)$ - easier to think about mechanistically. In fact, $\Lambda(t_2, s) - \Lambda(t_1, s)$ provides the intensity for the interval $(t_1, t_2]$.
- Dynamics in $\Lambda(t, s)$ provide dynamics for the discretized spatial point process

Illustrative growth models (each of which has an explicit solution)

- Exponential growth

$$\frac{d\Lambda(t,s)}{dt} = r(s)\,\Lambda(t,s)$$

- Gompertz growth

$$\frac{d\Lambda(t,s)}{dt} = r(s)e^{-\alpha(s)t}\Lambda(t,s)$$

- Logistic growth

$$\frac{d\Lambda(t,s)}{dt} = r(s)\Lambda(t,s)\left[1-\frac{\Lambda(t,s)}{K(s)}\right]$$

local growth rate

local carrying capacity

Process Models for the Parameters

$r(s),\ K(s)$ and initial intensity

$$\Lambda_0(s) = \int_{-\infty}^{0} \Omega(\tau, s)\, d\tau$$

are parameter processes which are modeled on log scale as

$$\log \Lambda_0(s) = \mu_\Lambda(s; \beta_\Lambda) + \theta_\Lambda(s), \quad \theta_\Lambda(s) \sim GP(0, C_\Lambda(\phi_\Lambda))$$

$$\log r(s) = \mu_r(s; \beta_r) + \theta_r(s), \quad \theta_r(s) \sim GP(0, C_r(\phi_r))$$

$$\log K(s) = \mu_K(s; \beta_K) + \theta_K(s), \quad \theta_K(s) \sim GP(0, C_K(\phi_K))$$

- Hence, given $r(s),\ K(s)$ and $\Lambda_0(s)$ the growth curve is fixed. Also, the μ’s are trend surfaces.

The problem

- We want the differential equations to be dependent at every location BUT
- We would not insist that the process exactly follows a logistic differential equation at every location
- We want to introduce some noise so convert to an SDE
- Again a time varying rate at each location
- Again $Z(t(s)) = \log r(t(s))$ with

$$dZ(t(s)) = \xi(Z(t(s)) - \log r(s)) + \sigma_Z dW(t,s)$$

- Induces a stationary process with a separable space-time covariance function

Discretizing Time (Euler Approximation)

Back to the original model, the intensity for the spatial point pattern in a time interval:

$$\Delta\Lambda_j(s) = \int_{t_{j,1}}^{t_{j,2}} \Omega(\tau, s)\, d\tau \approx \Omega(t_{j,1}, s)\,\Delta t$$

- Difference equation model:

$$\Delta\Lambda_j(s) = r(s)\Lambda_{j-1}(s)\left[1 - \frac{\Lambda_{j-1}(s)}{K(s)}\right]\Delta t$$

$$\Lambda_j(s) = \Lambda_0(s) + \sum_{l=1}^{j-1} \Delta\Lambda_l(s)$$

explicit transition

a recursion

Discrete-time Model

- Model parameters and latent processes:

$$\theta_r(s),\ \theta_K(s) \text{ and } \theta_\Lambda(s)$$
$$\beta_\Lambda, \beta_r, \beta_r$$
$$\phi_\Lambda, \phi_r, \phi_K$$

- Likelihood

point *i* in period *j*

$$\prod_{j=1}^{J}\left\{\exp\left(-\int_D \Delta\Lambda_j(s)\,ds\right)\prod_{i=1}^{n_j}\Delta\Lambda_j(x_{ji})\right\}\cdot\exp\left(-\int_D \Lambda_0(s)\,ds\right)\prod_{i=1}^{n_0}\Lambda_0(x_{0i})$$

stochastic integral

Discretizing Space

Divide region D into M cells. Rescaling and assuming homogeneous intensity in each cell. We obtain (with r(m), k(m) average growth rate and cumulative carrying capacity):

$$\frac{d\Lambda(t,m)}{dt} = r(m)\Lambda(t,m)[1 - \frac{\Lambda(t,m)}{K(m)}]$$

with induced transition

$$\Delta\Lambda_j(m) = r(m)\Lambda_{j-1}(m)\left[1 - \frac{\Lambda_{j-1}(m)}{K(m)}\right]\Delta t.$$

The joint likelihood (product Poisson):

$$\prod_{j=1}^{J}\left[\exp\left(-\sum_{m=1}^{M}\Delta\Lambda_j(m)\,A(m)\right)\prod_{m=1}^{M}\Delta\Lambda_j(m)^{n_{jm}}\right]$$

$$\cdot\exp\left(-\sum_{m=1}^{M}\Lambda_0(m)\,A(m)\right)\prod_{m=1}^{M}\Lambda_0(m)^{n_{0m}},$$