The SPDE/GMRF-based Approach to The Spatio-Temporal Dynamics of Crowdfunding Campaign

Han Yu, PhD

University of Northern Colorado

August 4, 2021

CBMS Conference, University of Alabama in Huntsville

Outline

1. Introduction
    - The Phenomenon of Crowdfunding
    - Research Paradigm
    - Methodology
        - Trajectory Modeling (TM)
        - SPDE/GMRF Approach
2. GMRF
3. SPDE-based GMRF
4. Application

New phenomena in entrepreneurship

- Crowdfunding Campaigns
  - One of the novelties of the emerging FinTech sector:

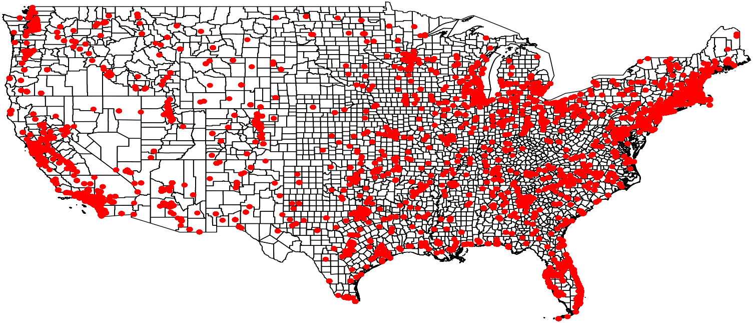
    Founders launch a campaign on an internet-based platform for a crowd of potential funders within limited time.
  - Potentially disruptive innovation for financing a variety of new entrepreneurial ventures without traditional financial intermediaries.
- Ways of Success
  - Identification: what are the key areas of success factors for crowdfunding campaigns
  - Explanation: how and why the percentage of a project's goal actually raised is dynamically associated with dollar pledged and funder count that reflect the signals of underlying project quality

Crowdfunding Campaign

- Over a billion dollars spent by millions of individual crowdfunders
- Large-scale action by the US Congress to encourage crowdfunding as a source of capital for new ventures.
- Both practice and policy continue to rapidly advance

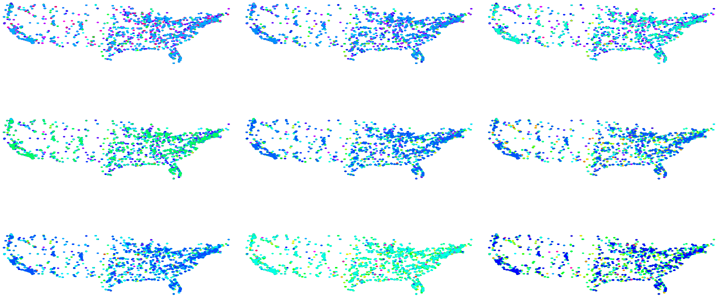
Distribution

- Little a priori basic academic knowledge of the distribution of crowdfunding mechanisms and of successful crowdfunding if any



Dynamics of Mechanisms

- No a priori knowledge of whether crowdfunding efforts reinforce or contradict existing theories about how ventures raise capital and achieve success either.
- Let alone read off
  - which variables are "important" in the dynamic process.
  - whether crowdfunding efforts reinforce or contradict existing theories about how ventures raise capital and achieve success.



Hierarchy of Claims

- Association
- Intervention
- Counterfactuals

Multivariate Identification and Interpretation

- The important and growing area of entrepreneurial activity and government action is understudied outside of the still-uncommon analysis of particular crowdfunding efforts (Mollick, 2014).
- The claim of independent association between one or more variables and outcome of interest requires adjustment for potentially confounding variables.
- Effort to adjust for confounding
  - Many variables are measured in .
    - The Kickstarter campaign of n=99,036 projects totaling about 1 billion USD in pledges from 2009 to 2017.
    - Raw covariates: *goal, pledged, deadline, created, launched, staff-pick, backers-count, deadline, goal, duration, latitude, longitude, time.*
  - A wide net is cast and multivariate models are build.
    - Technical covariates: high-dimensional

Existing Methods in The Literature

- As crowdfunding becomes more and more popular, many researchers have explored various methods to understand the phenomenon.

- More recently, researchers have employed various machine learning algorithms to the study

  - "black-box", "model-free", "model-blind", or "data-centric"

  - "function-fitting" (Darwiche, 2017): fitting data by a complex function defined by the neural network architecture.

Machine Learning To Causal Inference

- Despite failure stories (Shalev-Shwartz et al., 2017), albeit less publicized, the dramatic success in machine learning has led to an explosion of AI applications and increasing expectations for autonomous systems that exhibit human-level intelligence.
- The fundamental obstacles to the expectations
  - adaptability or robustness
  - explainability
  - the understanding of causal relationships, a necessary (though not sufficient) ingredient for achieving human-level intelligence
    - a parsimonious and modular representation of environment to answer interventional questions and counterfactual questions
    - statistical methods that have desirable statistical properties while remaining computationally feasible.
- Identification of factors influencing the success rate via SPDE/GMRF spatio-temporal approach.

Unmeasured Confounders

- Unmeasured confounding is a concern in many observational studies.
- Adjusting for unmeasured confounding permits us to estimate causal effects in nonexperimental studies.
- Traditional statistics is strong in devising ways of describing data and inferring distributional parameters from sample.
- Causal inference invokes non-parametric structural equations models as a formal and meaningful language for defining causal quantities, formulating causal assumptions, testing identifiability, and explicating many concepts used in causal discourse.

Coping with Unmeasured Confounding

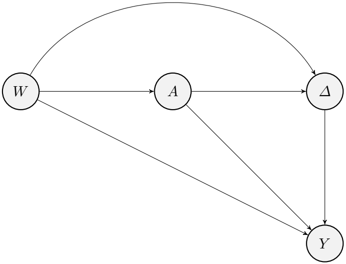
- Propensity score, regression and matching methods only control for measured confounders and do not control for unmeasured confounders.

- Approaches to unmeasured confounding control

  - Randomization control trials (RCTs)— interventional design

  - Instrumental variables (IVs)— observational study

Graphical Model



SEM

- SCM (Pearl, 2009)

$$
\begin{aligned}
W &= f_W(U_W), \\
A &= f_A(W, U_A), \\
\mathit{\Delta} &= f_{\mathit{\Delta}}(W, A, U_{\mathit{\Delta}}), \\
Y &= f_Y(W, A, \mathit{\Delta}, U_Y)
\end{aligned}
$$

- Counterfactuals Y^1 and Y^0 corresponding with interventions setting A and $\mathit{\Delta}$.
- ATE

$$
\begin{aligned}
\mathit{\Psi}(P) &= EY^1 - EY^0 \\
&= E\left[E(Y|do(A=1), \mathit{\Delta}=1, W) - E(Y|do(A=0), \mathit{\Delta}=1, W)\right]
\end{aligned}
$$

- The SCM is nonparametric SEM (NPSEM) when the realistic assumptions are involved.

Data Structures

- Simple data structure: O=(L, A, Y)$\sim P_0$ without common issues such as missingness and censoring.
    - O: realization of P_0
    - O_i: sample of P_0
    - o_i: observation of P_0

- Complex data structure
    - Censoring data structure: $O_{st} = (L, A, \widetilde{T}, \mathit{\Delta})_{st} \sim P_0$.
        - ★ T: time to event (deadline)
        - ★ C: censoring time (right censoring in the USA cohort)
        - ★ $\widetilde{T} = min(T, C)$ represents the T or C that was observed first
        - ★ $\mathit{\Delta}$ = I $(T \leq \widetilde{T})$ = I$(C \geq T)$: indicator that T was observed at or before C
    - Missingness data structure: $O = (L, A, \mathit{\Delta}, \mathit{\Delta}Y)_{st}$ with indicator $\mathit{\Delta}$ of missingness

Trajectory Modeling (TM)

- The evolution of an outcome of interest over time called developmental trajectory describes the progression of any behavioral, biological or physical phenomenon.

- Representing and understanding developmental trajectories is among the most fundamental and empirical important research topics in the social and behavioral sciences and medicine.

- To analyze the developmental trajectories,
  - Nagin and Land (1993) laid out the popular statistical method called group-based trajectory modeling (GBTM) to address issues related to the "hot topic" of the time—the criminal career debate.
  - Yu et. al. (2021) proposed network-based trajectory modeling (NBTM) as an extension of GBTM.

STP-NBTM

- STP-NBTM is obtained by assigning Gaussian priors to all elements of the latent field that can answer complex longitudinal questions
  - using heterogeneity of structural longitudinal (big, complex, and dynamic) data
  - incorporating targeted ensemble machine learning algorithms (super learner) to estimate quantity of interest while still maintain strong theoretical foundation providing valid inference
  - avoiding reliance on human art and unrealistic parametric models
- STP-NBTM can be expressed as a hierarchical model

Hierarchical Representation

- NBTM is obtained by assigning Gaussian priors to all elements of the latent field $\boldsymbol{x}$.

- The NBTM can schematically be represented as a hierarchical model (HM). For $i = 1, 2, \cdots, n$,

$$\begin{aligned}
\text{Data Model}: \quad & Y_i(t) | \eta_i(t), \boldsymbol{\theta}_d \sim \mathcal{D}(\eta_i(t), \boldsymbol{\theta}_d), \\
\text{Latent Random Field}: \quad & \eta_i(t) = \beta_0 + \sum_{m=1}^{p} \beta_m x_{mi}(t) + \omega_i(t), \\
\text{Dynamic Process}: \quad & \omega_i(t) = \mathcal{M}_\tau(\boldsymbol{\omega}(t), \boldsymbol{\Phi}_\omega) + \xi_i(t) \\
\text{Regularization}: \quad & \boldsymbol{\Phi}_\omega \sim \pi_1(\boldsymbol{\theta}_r), \\
\text{Residual Process}: \quad & \xi_i(t) \sim \pi_2(\boldsymbol{\theta}_u), \\
\text{Parameters}: \quad & \boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}, \boldsymbol{\theta}_d, \boldsymbol{\theta}_r, \boldsymbol{\theta}_u) \sim \pi(\beta_0, \boldsymbol{\beta}, \boldsymbol{\theta}_d, \boldsymbol{\theta}_r, \boldsymbol{\theta}_u).
\end{aligned}$$

Inference Framework — three-stage hierarchical model

- observations ($\boldsymbol{y}$):
  - conditionally independent given $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$

$$\boldsymbol{y}|\boldsymbol{\eta},\boldsymbol{\theta} \sim \prod_i p(y_i|\eta_i,\boldsymbol{\theta}).$$

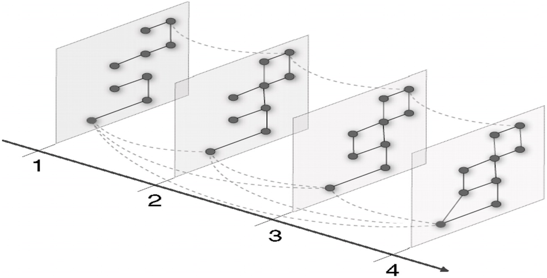
- latent field ($\boldsymbol{x}$):
  - a Gaussian Markov Random Field (GMRF) with sparse precision matrix $\boldsymbol{Q(\theta)}$

$$\boldsymbol{x}|\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0},\boldsymbol{Q}^{-1}(\boldsymbol{\theta})).$$

- hyperparameters ($\boldsymbol{\theta}$):
  - Precision parameters of the Gaussian priors assigned to latent field

$$\boldsymbol{\theta} \sim \pi(\boldsymbol{\theta}).$$

STP

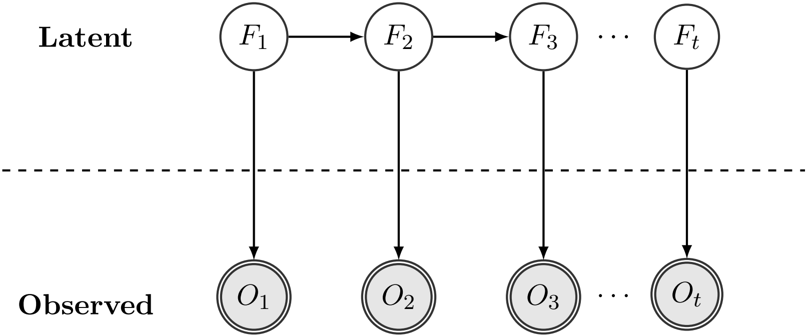


- The random variable $Z_{\boldsymbol{s}t}$, $\boldsymbol{s} \in D_t$ and $t \in T$, where $D_t = D(t)$ representing an evolution of random sets.

- The spatial locations in D_t are linked by the SPDE model object at $\forall t \in T$, while the process evolves, for example, according to an RW process across time.

Markov Random Field

- Outcomes $\boldsymbol{O} = (O_1, \cdots, O_t, \cdots)$ observed repeatedly from the state of a system characterized by a Latent MRF $\boldsymbol{F} = (F_1, F_2, \cdots, F_t, \cdots)$.



Random Fields

- Probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Y-valued random field (RF) is a collection of Y-valued random variables indexed by elements in a topological space D_t, i.e.,
  - a random field $\boldsymbol{Y}$ is a collection
  $$\{Y_{st}(\omega) : s \in D_t, t \in T, \omega \in \Omega\}$$
  - probabilistic model: product measure on product space.

- RFs confront astronomers, physicists, geologists, meteorologists, biologists, and other natural scientists.

- RFs even underlie the processes of social and economic change.

Gaussian Fields

- The process $\{z(\boldsymbol{s}), \boldsymbol{s} \in D\}$ is a Gaussian field if for any $k \geq 1$ and any locations $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_k \in D$, $(\boldsymbol{z}(\boldsymbol{s}_1), \cdots, \boldsymbol{z}(\boldsymbol{s}_k))^T$ is normally distributed. The mean function and covariance function (CF) of $\boldsymbol{z}$ are
$$\mu(\boldsymbol{s}) = E(\boldsymbol{z}(\boldsymbol{s})),\ C(\boldsymbol{s}, \boldsymbol{t}) = Cov(\boldsymbol{z}(\boldsymbol{s}), \boldsymbol{z}(\boldsymbol{t})),$$
which are both assumed to exist for all $\boldsymbol{s}$ and $\boldsymbol{t}$.

- In the machine learning literature, the phrase "Gaussian process models" is often used (Rasmussen and Williams, 2006).

- By modern definitions, a RF is a generalization of a stochastic process where the underlying parameter need no longer be real or integer valued "time" but can instead take values that are multidimensional vectors or points on some manifold (Vanmarcke, 2010).

Covariance Structure: Covariance Functions

- The random effects are captured through a spatial covariance model, especially in GF.
- In most applications, one of the following isotropic CFs is used in geostatistics:

Exponential	$C(h) = exp(-3h)$
Gaussian	$C(h) = exp(-3h^2)$
Powered exponential	$C(h) = exp(-3h^\alpha),\ 0 < \alpha \leq 2$
Matérn	$C(h) = \dfrac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(s_\nu h)^\nu K_\nu(s_\nu h).$

  - K_ν is the modified Bessel function of the second kind and order $\nu > 0$, and s_ν is a function of ν such that the covariance function is scaled to $C(1) = 0.05$.
  - For similar reasons, the multiplicative factor 3 enters in the exponent of the exponential, Gaussian and powered exponential CF, where now $C(1) = 0.04979 \approx 0.05$.

SPDE-based GMRF

- A large class of random field models can be expressed as solutions to continuous domain stochastic partial differential equations (SPDEs) (Lindgren et al., 2011, Simpson, Lindgren, and Rue, 2012a,b) with explicit links between the parameters of each SPDE and the elements of precision matrices for weights in a discrete basis function representation.
- Such models include those with Matérn covariance functions (Whittle, 1963).
- In contrast to covariance-based models it is far easier to introduce non-stationarity into the SPDE models because the differential operators act locally and only mild regularity conditions are required.
- Classical Gaussian random fields can be merged with methods based on the Markov property, providing continuous domain models that are computationally efficient, and where the parameters can be specified locally without having to worry about positive definiteness of covariance functions.

Fitting GMRFs to GFs

- The (continuous) stationary Matérn fields fields are derived from SPDEs
$$(\kappa^2 - \mathit{\Delta})^{\alpha/2} x(\boldsymbol{s})) = \mathcal{W}(\boldsymbol{s}),\ \ \boldsymbol{s} \in D^d$$
- The (continuous) non-stationary Gaussian fields are derived from SPDEs
$$(\kappa(\boldsymbol{s})^2 - \mathit{\Delta})^{\alpha/2} (\tau(\boldsymbol{s}) x(\boldsymbol{s})) = \mathcal{W}(\boldsymbol{s}),\ \ \boldsymbol{s} \in D^d$$
where $\mathit{\Delta}$ is the Laplacian, $\kappa > 0$ is the spatial scale parameter, α controls the smoothness, τ controls the variance, and $W(\boldsymbol{s})$ is a Gaussian spatial white noise processes.
- The SPDE model is defined with considering the PC-prior derived in Fuglstad et al. (2018).
- The solution is a Gaussian field with Matérn covariance function having smoothness $\nu = \alpha - \frac{d}{2}$.

Kronecker Product Model

- The most important method is to construct a Kronecker product model, starting from a basis representation

$$x(\boldsymbol{s},t) = \sum_k \psi_k(\boldsymbol{s},t) x_k,$$

- With each basis function is the product of a spatial and a temporal basis function, $\psi_k(\boldsymbol{s},t) = \psi_i^{\boldsymbol{s}}(\boldsymbol{s})\psi_j^t(t)$, the space-time SPDE

$$\frac{\partial}{\partial t}(\kappa(\boldsymbol{s})^2 - \Delta)^{\alpha/2}(\tau(\boldsymbol{s})x(\boldsymbol{s},t)) = \mathcal{W}(\boldsymbol{s},t),\ (\boldsymbol{s},t) \in D \times \mathscr{R}$$

  generates a precision matrix for the weight vector $\boldsymbol{x}$ as $Q = Q_t \otimes Q_s$, where Q_s is the precision for the previous purely spatial model and Q_t is the precision corresponding to a one-dimensional random walk.

Results

Fixed effects:

	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$	mode
Goal	-0.072	0.001	-0.074	-0.072	-0.069	-0.072
Pledged	0.095	0.001	0.092	0.095	0.098	0.095

Random effects:

	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$	mode
Range r	139.247	29.162	92.428	135.604	206.523	128.277
σ	0.141	0.042	0.077	0.134	0.241	0.122
ρ	0.962	0.029	0.884	0.970	0.993	0.982