# Estimación de la intensidad de fondo en procesos de Hawkes mediante transporte óptimo

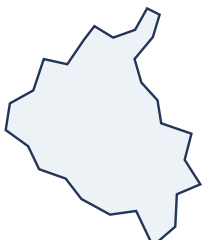
Arturo Jaramillo Gil

Centro de Investigación en Matemáticas (CIMAT)

Trabajo conjunto con Isaías Bañales

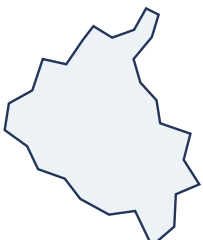
# Dónde ocurren los sismos?

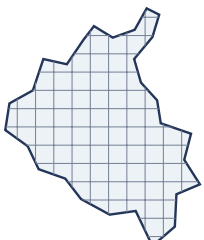
**Ley de eventos raros como punto de inicio.**



# Dónde ocurren los sismos?

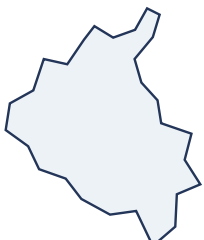
**Ley de eventos raros como punto de inicio.**

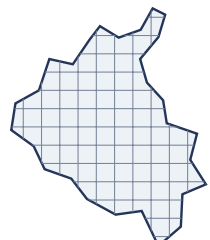




# Dónde ocurren los sismos?

**Ley de eventos raros como punto de inicio.**

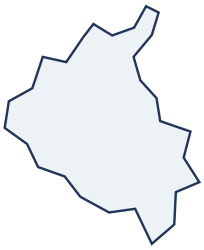




Usamos entonces la aproximación Poisson a la binomial.

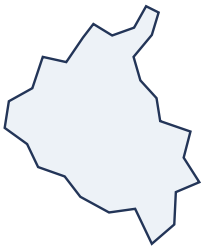
# Modelación Poisson para sismos

Poisson para una región

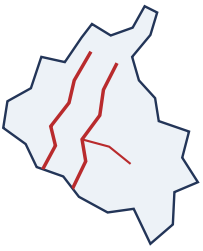


# Modelación Poisson para sismos

Poisson para una región

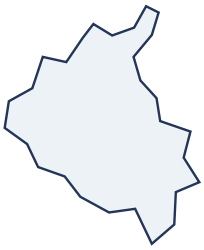


Que tal si la región tiene fallas?

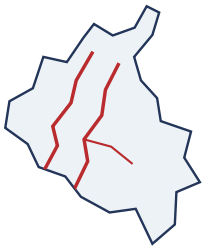


# Modelación Poisson para sismos

Poisson para una región



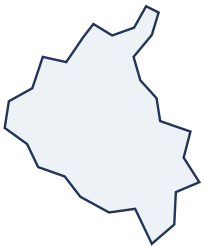
Que tal si la región tiene fallas?



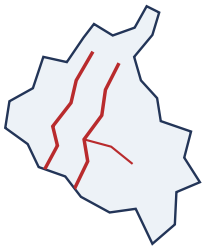
Consideramos Poisson no homogéneos con tasa $\rho(x)$.

# Modelación Poisson para sismos

Poisson para una región



Que tal si la región tiene fallas?



Consideramos Poisson no homogéneos con tasa $\rho(x)$.

**Pero hay réplicas...**

# Réplicas y autoexcitación

Una posible modelación es tomar tasas

$$\lambda_\rho(t, x) = \rho(x) + B_t(x),$$

donde $\rho(x)$ es la intensidad de fondo y

$$B_t(x) = \sum_{t_i < t} \nu(t - t_i, x - x_i, r_i)$$

es la excitación acumulada por los eventos anteriores.

# Réplicas y autoexcitación

Una posible modelación es tomar tasas

$$\lambda_\rho(t, x) = \rho(x) + B_t(x),$$

donde $\rho(x)$ es la intensidad de fondo y

$$B_t(x) = \sum_{t_i < t} \nu(t - t_i, x - x_i, r_i)$$

es la excitación acumulada por los eventos anteriores. Esto define un proceso de conteo llamado proceso de Hawkes.

# Réplicas y autoexcitación

Una posible modelación es tomar tasas

$$\lambda_\rho(t, x) = \rho(x) + B_t(x),$$

donde $\rho(x)$ es la intensidad de fondo y

$$B_t(x) = \sum_{t_i < t} \nu(t - t_i, x - x_i, r_i)$$

es la excitación acumulada por los eventos anteriores. Esto define un proceso de conteo llamado proceso de Hawkes. En el estudio de sismos, se dan formas concretas de $\nu$, y el modelo resultante se conoce como ETAS.

# Réplicas y autoexcitación

Una posible modelación es tomar tasas

$$\lambda_\rho(t, x) = \rho(x) + B_t(x),$$

donde $\rho(x)$ es la intensidad de fondo y

$$B_t(x) = \sum_{t_i < t} \nu(t - t_i, x - x_i, r_i)$$

es la excitación acumulada por los eventos anteriores. Esto define un proceso de conteo llamado proceso de Hawkes. En el estudio de sismos, se dan formas concretas de $\nu$, y el modelo resultante se conoce como ETAS.

**Objetivo**

Como obtener información sobre $\rho$, si ya conozco $\nu$.

# Verosimilitud

- Observamos datos $(t_q, x_q, r_q)_{q \geq 1}$ en $[0, T]$, con $t_q$ tiempos, $x_q$ posiciones y $r_q$ intensidades.

# Verosimilitud

- Observamos datos $(t_q, x_q, r_q)_{q \geq 1}$ en $[0, T]$, con $t_q$ tiempos, $x_q$ posiciones y $r_q$ intensidades.
- El total de observaciones es una variable aleatoria $N_T$.

# Verosimilitud

- Observamos datos $(t_q, x_q, r_q)_{q\geq 1}$ en $[0, T]$, con $t_q$ tiempos, $x_q$ posiciones y $r_q$ intensidades.
- El total de observaciones es una variable aleatoria $N_T$.
- La log-verosimilitud normalizada, salvo términos independientes de $\rho$, es:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

# Verosimilitud

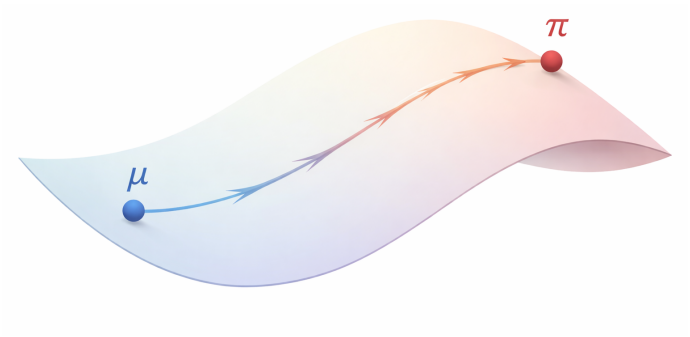
- Observamos datos $(t_q, x_q, r_q)_{q\geq 1}$ en $[0, T]$, con $t_q$ tiempos, $x_q$ posiciones y $r_q$ intensidades.
- El total de observaciones es una variable aleatoria $N_T$.
- La log-verosimilitud normalizada, salvo términos independientes de $\rho$, es:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

¿Podemos intentar maximizar con respecto a $\rho$?

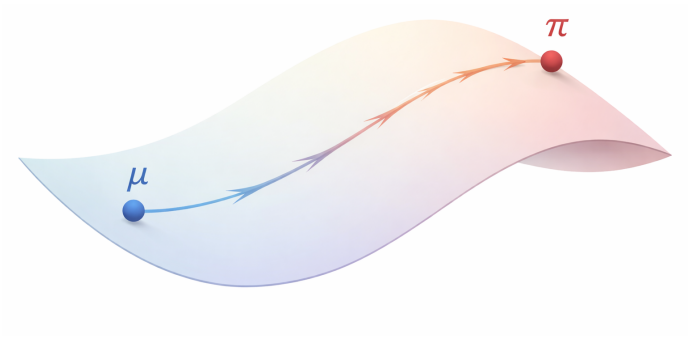
# Perspectiva variacional

Mover una medida $\mu$ para aumentar una función $\mathcal{F}$, llegando a un límite $\pi$.



# Perspectiva variacional

Mover una medida $\mu$ para aumentar una función $\mathcal{F}$, llegando a un límite $\pi$.




Tomaremos ideas de las superficies euclidianas.

# Gradientes y geometría

En $\mathbb{R}^m$, para maximizar una función $F$:

$$\dot{\theta}_u = \nabla F(\theta_u), \qquad \frac{d}{du} F(\theta_u) = |\nabla F(\theta_u)|^2$$

# Gradientes y geometría

En $\mathbb{R}^m$, para maximizar una función $F$:

$$\dot{\theta}_u = \nabla F(\theta_u), \qquad \frac{d}{du} F(\theta_u) = \text{magnitud}(\text{dirección destacada}[F](\theta_u))$$

# Gradientes y geometría

En $\mathbb{R}^m$, para maximizar una función $F$:

$$\dot{\theta}_u = \nabla F(\theta_u), \qquad \frac{d}{du}F(\theta_u) = \text{magnitud}(\text{dirección destacada}[F](\theta_u))$$

Si el input son medidas, elegimos una métrica estilo Riemanniano **(nota)** que sustituya a 'magnitud', así como un espacio con una dirección destacada

$$\dot{\mu}_u = \mathbb{W}_g \mathcal{F}(\mu_u), \qquad \frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_g \mathcal{F}(\mu_u)\|^2_{g,\mu_u}.$$

# Gradientes y geometría

En $\mathbb{R}^m$, para maximizar una función $F$:

$$\dot{\theta}_u = \nabla F(\theta_u), \qquad \frac{d}{du}F(\theta_u) = \text{magnitud}(\text{dirección destacada}[F](\theta_u))$$

Si el input son medidas, elegimos una métrica estilo Riemanniano **(nota)** que sustituya a 'magnitud', así como un espacio con una dirección destacada

$$\dot{\mu}_u = \mathbb{W}_g \mathcal{F}(\mu_u), \qquad \frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_g \mathcal{F}(\mu_u)\|^2_{g,\mu_u}.$$

La $g$ alude a la geometría que elegimos. Cada geometría debe determinar una dirección de ascenso y su magnitud.

# Geometría de Wasserstein

La primera variación $\delta\mathcal{F}/\delta\rho : \mathbb{R}^d \to \mathbb{R}$ es una función t.q.

$$\frac{d}{d\varepsilon}\mathcal{F}((\rho + \varepsilon h)\, dx)\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d} \delta\mathcal{F}/\delta\rho(x) h(x)\, dx.$$

# Geometría de Wasserstein

La primera variación $\delta\mathcal{F}/\delta\rho : \mathbb{R}^d \to \mathbb{R}$ es una función t.q.

$$\frac{d}{d\varepsilon}\mathcal{F}((\rho + \varepsilon h)\, dx)\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d} \delta\mathcal{F}/\delta\rho(x) h(x)\, dx.$$

**Caja negra:** una propuesta es proponer

$$\begin{aligned} \mathbb{W}_W\mathcal{F}(\mu) &= -\operatorname{div}(\rho \nabla \delta\mathcal{F}/\delta\rho) \\ \|v\|^2_{W,\mu} &= \inf_{-\operatorname{div}(\rho w)=v} \int_{\mathbb{R}^d} |w(x)|^2\, \mu(dx). \end{aligned}$$

# Geometría de Wasserstein

La primera variación $\delta\mathcal{F}/\delta\rho : \mathbb{R}^d \to \mathbb{R}$ es una función t.q.

$$\frac{d}{d\varepsilon}\mathcal{F}((\rho + \varepsilon h)\, dx)\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d} \delta\mathcal{F}/\delta\rho(x) h(x)\, dx.$$

**Caja negra:** una propuesta es proponer

$$\mathbb{W}_W\mathcal{F}(\mu) = -\operatorname{div}(\rho\nabla\delta\mathcal{F}/\delta\rho)$$
$$\|v\|^2_{W,\mu} = \inf_{-\operatorname{div}(\rho w)=v} \int_{\mathbb{R}^d} |w(x)|^2\, \mu(dx).$$

Puede encontrarse una trayectoria de medidas $\mu_u$, t.q.

$$\frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_W\mathcal{F}(\mu_u)\|^2_{W,\mu_u}$$

**(nota)**

# Geometría de Fisher-Rao

**Caja negra:** una propuesta alternativa es

$$\nabla_{FR}\mathcal{F}(\mu) = \rho\,\frac{\delta \mathcal{F}}{\delta \rho},$$

$$\|v\|^2_{FR,\mu} = \inf_{\rho a = v} \int_{\mathbb{R}^d} |a(x)|^2\,\mu(dx).$$

# Geometría de Fisher-Rao

**Caja negra:** una propuesta alternativa es

$$\mathbb{W}_{FR}\mathcal{F}(\mu) = \rho\,\frac{\delta\mathcal{F}}{\delta\rho},$$

$$\|v\|^2_{FR,\mu} = \inf_{\rho a = v} \int_{\mathbb{R}^d} |a(x)|^2\,\mu(dx).$$

Puede encontrarse una trayectoria de medidas $\mu_u$, t.q.

$$\frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_{FR}\mathcal{F}(\mu_u)\|^2_{FR,\mu_u}.$$

**(nota)**

# Geometría de Wasserstein-Fisher-Rao

Combinamos las dos propuestas, con $\kappa_W, \kappa_{FR} > 0$:

$$\mathbb{W}_{WFR}\mathcal{F}(\mu) = -\kappa_W \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \kappa_{FR}\rho \frac{\delta \mathcal{F}}{\delta \rho},$$

$$\|v\|^2_{WFR,\mu} = \inf_{-\operatorname{div}(\rho w)+\rho a = v} \int_{\mathbb{R}^d} \left( \frac{|w(x)|^2}{\kappa_W} + \frac{a(x)^2}{\kappa_{FR}} \right) \mu(dx).$$

# Geometría de Wasserstein-Fisher-Rao

Combinamos las dos propuestas, con $\kappa_W, \kappa_{FR} > 0$:

$$\mathbb{W}_{WFR}\mathcal{F}(\mu) = -\kappa_W \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \kappa_{FR}\rho \frac{\delta \mathcal{F}}{\delta \rho},$$

$$\|v\|^2_{WFR,\mu} = \inf_{-\operatorname{div}(\rho w)+\rho a = v} \int_{\mathbb{R}^d} \left( \frac{|w(x)|^2}{\kappa_W} + \frac{a(x)^2}{\kappa_{FR}} \right) \mu(dx).$$

Puede encontrarse una trayectoria de medidas $\mu_u$, t.q.

$$\frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_{WFR}\mathcal{F}(\mu_u)\|^2_{WFR,\mu_u}.$$

# Geometría de Wasserstein-Fisher-Rao

Combinamos las dos propuestas, con $\kappa_W, \kappa_{FR} > 0$:

$$\mathbb{W}_{WFR}\mathcal{F}(\mu) = -\kappa_W \operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \kappa_{FR}\rho \frac{\delta \mathcal{F}}{\delta \rho},$$

$$\|v\|^2_{WFR,\mu} = \inf_{-\operatorname{div}(\rho w)+\rho a = v} \int_{\mathbb{R}^d} \left( \frac{|w(x)|^2}{\kappa_W} + \frac{a(x)^2}{\kappa_{FR}} \right) \mu(dx).$$

Puede encontrarse una trayectoria de medidas $\mu_u$, t.q.

$$\frac{d}{du}\mathcal{F}(\mu_u) = \|\mathbb{W}_{WFR}\mathcal{F}(\mu_u)\|^2_{WFR,\mu_u}.$$

**(nota)**

# En nuestra aplicación esto no tiene sentido

Recordemos la log-verosimilitud:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

# En nuestra aplicación esto no tiene sentido

Recordemos la log-verosimilitud:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

Su primera variación es

$$\frac{\delta \mathcal{F}_T}{\delta \rho} = \frac{1}{T}\sum_{q=1}^{N_T} \frac{\delta_{x_q}}{\rho(x_q) + B_{t_q}(x_q)} - 1.$$

# En nuestra aplicación esto no tiene sentido

Recordemos la log-verosimilitud:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

Su primera variación es

$$\frac{\delta \mathcal{F}_T}{\delta \rho} = \frac{1}{T}\sum_{q=1}^{N_T} \frac{\delta_{x_q}}{\rho(x_q) + B_{t_q}(x_q)} - 1.$$

Las deltas de Dirac dan variación distribucional...

# En nuestra aplicación esto no tiene sentido

Recordemos la log-verosimilitud:

$$\mathcal{F}_T(\rho) = \frac{1}{T}\sum_{q=1}^{N_T} \log\left(\rho(x_q) + B_{t_q}(x_q)\right) - \int_{\mathbb{R}^d} \rho(x)\, dx.$$

Su primera variación es

$$\frac{\delta \mathcal{F}_T}{\delta \rho} = \frac{1}{T}\sum_{q=1}^{N_T} \frac{\delta_{x_q}}{\rho(x_q) + B_{t_q}(x_q)} - 1.$$

Las deltas de Dirac dan variación distribucional... pero estabamos muy cerca :-(

# Proyección sobre funciones

Observemos que si $\phi$ es función de prueba, se tiene que

$$\langle \mathbb{W}_{WFR}\mathcal{F}_T(\rho), \phi\rangle = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho\nabla\phi)(x_q) + \kappa_{FR}\rho(x_q)\phi(x_q)}{\rho(x_q) + B_{t_q}(x_q)} - \kappa_{FR}\int_{\mathbb{R}^d} \rho(x)\phi(x)\,dx.$$

# Proyección sobre funciones

Observemos que si $\phi$ es función de prueba, se tiene que

$$\langle \mathbb{W}_{WFR}\mathcal{F}_T(\rho), \phi \rangle = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho\nabla\phi)(x_q) + \kappa_{FR}\rho(x_q)\phi(x_q)}{\rho(x_q) + B_{t_q}(x_q)} - \kappa_{FR}\int_{\mathbb{R}^d} \rho(x)\phi(x)\,dx.$$

El lado izquierdo se puede ver como una proyección en $L^2$.

# Proyección sobre funciones

Observemos que si $\phi$ es función de prueba, se tiene que

$$\langle \mathbb{W}_{WFR}\mathcal{F}_T(\rho), \phi \rangle = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho\nabla\phi)(x_q) + \kappa_{FR}\rho(x_q)\phi(x_q)}{\rho(x_q) + B_{t_q}(x_q)} - \kappa_{FR}\int_{\mathbb{R}^d} \rho(x)\phi(x)\,dx.$$

El lado izquierdo se puede ver como una proyección en $L^2$.

**Estrategia**

- Proponer medidas con densidad parametrizada.

# Proyección sobre funciones

Observemos que si $\phi$ es función de prueba, se tiene que

$$\langle \mathbb{W}_{WFR}\mathcal{F}_T(\rho), \phi\rangle = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho\nabla\phi)(x_q) + \kappa_{FR}\rho(x_q)\phi(x_q)}{\rho(x_q) + B_{t_q}(x_q)} - \kappa_{FR}\int_{\mathbb{R}^d} \rho(x)\phi(x)\,dx.$$

El lado izquierdo se puede ver como una proyección en $L^2$.

**Estrategia**

- Proponer medidas con densidad parametrizada.
- Ver a la familia como una variedad.

# Proyección sobre funciones

Observemos que si $\phi$ es función de prueba, se tiene que

$$\langle \mathbb{W}_{WFR}\mathcal{F}_T(\rho), \phi \rangle = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho\nabla\phi)(x_q) + \kappa_{FR}\rho(x_q)\phi(x_q)}{\rho(x_q) + B_{t_q}(x_q)} - \kappa_{FR}\int_{\mathbb{R}^d} \rho(x)\phi(x)\,dx.$$

El lado izquierdo se puede ver como una proyección en $L^2$.

**Estrategia**

- Proponer medidas con densidad parametrizada.
- Ver a la familia como una variedad.
- Proyectar $\mathbb{W}_{WFR}\mathcal{F}_T(\rho)$ sobre esos espacios tangentes.

# Mezclas gaussianas

Sea $\chi_\sigma(x; m)$ la densidad normal media $m$ varianza $\sigma$. Consideremos la familia

$$\rho_\theta(x) = \sum_{i=1}^{k} e^{\alpha_i} \chi_\sigma(x; m_i), \qquad \theta = (\alpha_1, \ldots, \alpha_k, m_1, \ldots, m_k).$$

# Mezclas gaussianas

Sea $\chi_\sigma(x; m)$ la densidad normal media $m$ varianza $\sigma$. Consideremos la familia

$$\rho_\theta(x) = \sum_{i=1}^{k} e^{\alpha_i} \chi_\sigma(x; m_i), \qquad \theta = (\alpha_1, \ldots, \alpha_k, m_1, \ldots, m_k).$$

Su plano tangente está generado por las derivadas respecto de los parámetros:

$$d_j^\theta = \partial_{\theta_j} \rho_\theta, \qquad T_{\rho_\theta} = \operatorname{span}\{d_j^\theta\}.$$

# Mezclas gaussianas

Sea $\chi_\sigma(x; m)$ la densidad normal media $m$ varianza $\sigma$. Consideremos la familia

$$\rho_\theta(x) = \sum_{i=1}^{k} e^{\alpha_i} \chi_\sigma(x; m_i), \qquad \theta = (\alpha_1, \ldots, \alpha_k, m_1, \ldots, m_k).$$

Su plano tangente está generado por las derivadas respecto de los parámetros:

$$d_j^\theta = \partial_{\theta_j} \rho_\theta, \qquad T_{\rho_\theta} = \operatorname{span}\{d_j^\theta\}.$$

- Describimos formalmente la evolución de las densidades bajo el flujo WFR.

# Mezclas gaussianas

Sea $\chi_\sigma(x; m)$ la densidad normal media $m$ varianza $\sigma$. Consideremos la familia

$$\rho_\theta(x) = \sum_{i=1}^{k} e^{\alpha_i} \chi_\sigma(x; m_i), \qquad \theta = (\alpha_1, \ldots, \alpha_k, m_1, \ldots, m_k).$$

Su plano tangente está generado por las derivadas respecto de los parámetros:

$$d_j^\theta = \partial_{\theta_j} \rho_\theta, \qquad T_{\rho_\theta} = \operatorname{span}\{d_j^\theta\}.$$

- Describimos formalmente la evolución de las densidades bajo el flujo WFR.
- Expresamos la evolución de una mezcla gaussiana mediante las velocidades de sus parámetros.

# Mezclas gaussianas

Sea $\chi_\sigma(x; m)$ la densidad normal media $m$ varianza $\sigma$. Consideremos la familia

$$\rho_\theta(x) = \sum_{i=1}^{k} e^{\alpha_i} \chi_\sigma(x; m_i), \qquad \theta = (\alpha_1, \ldots, \alpha_k, m_1, \ldots, m_k).$$

Su plano tangente está generado por las derivadas respecto de los parámetros:

$$d_j^\theta = \partial_{\theta_j} \rho_\theta, \qquad T_{\rho_\theta} = \operatorname{span}\{d_j^\theta\}.$$

- Describimos formalmente la evolución de las densidades bajo el flujo WFR.
- Expresamos la evolución de una mezcla gaussiana mediante las velocidades de sus parámetros.
- Emparejamos ambas velocidades con los generadores del plano tangente e igualamos: obtenemos una dinámica para los parámetros.

# Cálculos explícitos

Definimos

$$G_{ij}(\theta) := \int_{\mathbb{R}^d} d_i^\theta(x) d_j^\theta(x)\, dx.$$

$$(g_T)_i(\theta) = \frac{1}{T} \sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho_\theta \nabla d_i^\theta)(x_q) + \kappa_{FR} \rho_\theta(x_q) d_i^\theta(x_q)}{\rho_\theta(x_q) + B_{t_q}(x_q)}$$
$$- \kappa_{FR} \int_{\mathbb{R}^d} \rho_\theta(x) d_i^\theta(x)\, dx.$$

# Cálculos explícitos

Definimos

$$G_{ij}(\theta) := \int_{\mathbb{R}^d} d_i^\theta(x) d_j^\theta(x)\, dx.$$

$$(g_T)_i(\theta) = \frac{1}{T} \sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho_\theta \nabla d_i^\theta)(x_q) + \kappa_{FR} \rho_\theta(x_q) d_i^\theta(x_q)}{\rho_\theta(x_q) + B_{t_q}(x_q)}$$

$$- \kappa_{FR} \int_{\mathbb{R}^d} \rho_\theta(x) d_i^\theta(x)\, dx.$$

El programa descrito en el slide anterior induce la dinámica

$$G(\theta(u))\dot{\theta}(u) = g_T(\theta(u)).$$

# Cálculos explícitos

Definimos

$$G_{ij}(\theta) := \int_{\mathbb{R}^d} d_i^\theta(x) d_j^\theta(x)\, dx.$$

$$(g_T)_i(\theta) = \frac{1}{T}\sum_{q=1}^{N_T} \frac{-\kappa_W \operatorname{div}(\rho_\theta \nabla d_i^\theta)(x_q) + \kappa_{FR}\rho_\theta(x_q) d_i^\theta(x_q)}{\rho_\theta(x_q) + B_{t_q}(x_q)}$$

$$- \kappa_{FR} \int_{\mathbb{R}^d} \rho_\theta(x) d_i^\theta(x)\, dx.$$

El programa descrito en el slide anterior induce la dinámica

$$G(\theta(u))\dot{\theta}(u) = g_T(\theta(u)).$$

Si resolvemos la ODE, nuestro candidato de aproximación sería

$$\rho_{\theta_\infty}(x) = \sum_{i=1}^{k} e^{\alpha_{i,\infty}} \chi_\sigma(x; m_{i,\infty}).$$

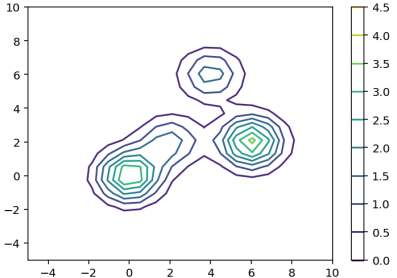
# Poisson no homogéneo

Un proceso de Poisson no homogéneo es un Hawkes sin memoria.

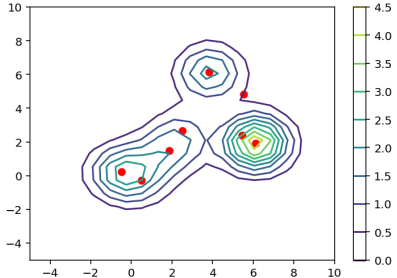
# Poisson no homogéneo

Un proceso de Poisson no homogéneo es un Hawkes sin memoria. Simulamos 695 eventos con una intensidad conocida y usamos sus posiciones para estimarla.

Intensidad verdadera



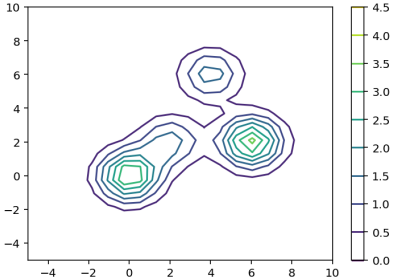

Intensidad estimada



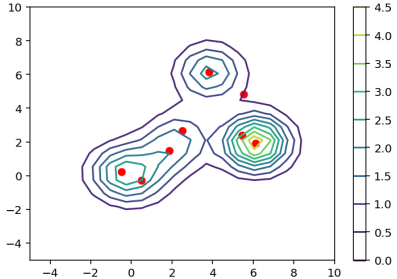
# Poisson no homogéneo

Un proceso de Poisson no homogéneo es un Hawkes sin memoria. Simulamos 695 eventos con una intensidad conocida y usamos sus posiciones para estimarla.

Intensidad verdadera




Intensidad estimada




De 100 componentes iniciales a 8 componentes finales.

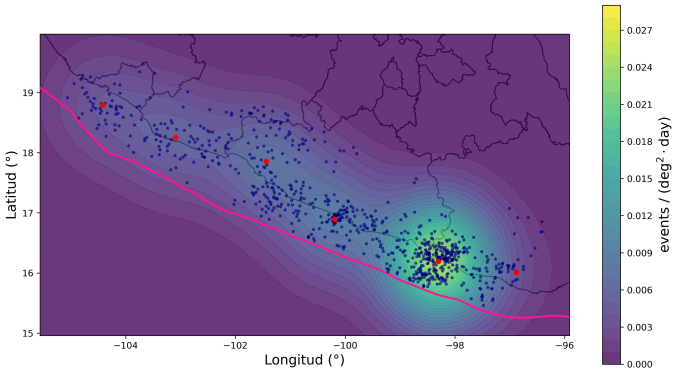
# Datos reales: sismos en la costa de México

Usamos un catálogo de sismos entre **2000 y 2016** para estimar la intensidad de fondo en el sur de México.



# Intensidad de fondo estimada

El modelo de Hawkes incorpora las réplicas y permite estimar la intensidad de fondo.




El fondo estimado se representa con solo 6 componentes gaussianas.

Los puntos rojos indican sus centros.

# Otras contribuciones

- Dar sentido al gradiente Wasserstein–Fisher–Rao cuando la primera variación es una distribución.

# Otras contribuciones

- Dar sentido al gradiente Wasserstein–Fisher–Rao cuando la primera variación es una distribución.
- Obtener una dinámica concreta mediante proyecciones sobre familias de densidades.

# Otras contribuciones

- Dar sentido al gradiente Wasserstein–Fisher–Rao cuando la primera variación es una distribución.
- Obtener una dinámica concreta mediante proyecciones sobre familias de densidades.
- Identificar una ecuación poblacional cuyo único cero es el fondo verdadero.

# Otras contribuciones

- Dar sentido al gradiente Wasserstein–Fisher–Rao cuando la primera variación es una distribución.
- Obtener una dinámica concreta mediante proyecciones sobre familias de densidades.
- Identificar una ecuación poblacional cuyo único cero es el fondo verdadero.
- Establecer consistencia condicional para los estimadores que satisfacen las ecuaciones de equilibrio proyectadas.

# Trabajo futuro

- Permitir una intensidad de fondo que cambie con el tiempo.

# Trabajo futuro

- Permitir una intensidad de fondo que cambie con el tiempo.
- Analizar la convergencia del esquema numérico y obtener tasas explícitas.

# Trabajo futuro

- Permitir una intensidad de fondo que cambie con el tiempo.
- Analizar la convergencia del esquema numérico y obtener tasas explícitas.
- Entender el efecto de la fusión de componentes gaussianas.

# Trabajo futuro

- Permitir una intensidad de fondo que cambie con el tiempo.
- Analizar la convergencia del esquema numérico y obtener tasas explícitas.
- Entender el efecto de la fusión de componentes gaussianas.
- Estimar de manera no paramétrica el mecanismo de autoexcitación junto con el fondo.

Contacto jagil@cimat.mx

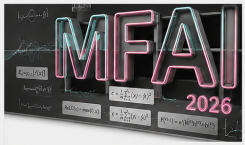



**Áreas de concentración**

Teoría de probabilidad, Teoría estadística, Modelación estocástica e inferencia en ciencias, industria, y tecnología, Probabilidad e inferencia estadística en ciencias de datos, Finanzas y riesgo

**Becas SECIHTI**

# Actividad en CIMAT




**20–23 de abril de 2026**

Mathematical Foundations of Artificial Intelligence, MFAI26

Auditorio IIMAS-UNAM

Informes: jcpardo@cimat.mx




**03–07 de agosto de 2026**

Latin American School of Information Theory and Statistics (LASITS 2026)

CIMAT, Guanajuato, México

https://sites.google.com/view/lasits-2026/home




**17–21 de agosto de 2026**

Tercer Taller Latinoamericano de Minería Genómica

https://bgc-latam.org/




**27 de agosto de 2026 • 05:00 pm**

Pláticas de estudiantes de Probabilidad y Estadística

Salón G001, CIMAT Guanajuato




**02 de septiembre de 2026 • 17:30 h**

Seminario de Finanzas:
Calificaciones ESG: Retos y Oportunidades

seminario.finanzasestocasticas@im.unam.mx

Muchas gracias!

# Bibliografía


[1] I. Bañales y A. Jaramillo.
*Projected Wasserstein-Fisher-Rao Flows for Hawkes Background-Intensity Estimation*.
arXiv:2610.19588, 2026.

[2] P. J. Laub, Y. Lee y T. Taimre.
*The Elements of Hawkes Processes*.
Springer, 2021.

[3] F. Santambrogio.
*Optimal Transport for Applied Mathematicians*.
Birkhäuser, 2015.

[4] S. Chewi, J. Niles-Weed y P. Rigollet.
*Statistical Optimal Transport*.
Springer, 2025.

[5] L. Chizat, G. Peyré, B. Schmitzer y F.-X. Vialard.
*An Interpolating Distance Between Optimal Transport and Fisher–Rao Metrics*.
Foundations of Computational Mathematics, 18:1–44, 2018.