

Resolución de problemas inversos utilizando métodos estadísticos

Karina Guadalupe Temperini

Universidad Nacional del Litoral
Instituto de Matemática Aplicada del Litoral

SEMINARIO CARLOS SEGOVIA FERNÁNDEZ

14 DE OCTUBRE DE 2016

Contexto

Problema Inverso: hallar x en

$$Tx = y \quad (P)$$

donde $T : \mathcal{X} \rightarrow \mathcal{Y}$ tal que $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, \mathcal{X}, \mathcal{Y} espacios de Hilbert; $\mathcal{R}(T)$ no cerrado; y : dato.

Contexto

Problema Inverso: hallar x en

$$Tx = y \quad (P)$$

donde $T : \mathcal{X} \rightarrow \mathcal{Y}$ tal que $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, \mathcal{X}, \mathcal{Y} espacios de Hilbert; $\mathcal{R}(T)$ no cerrado; y : dato.



El problema (P) es “mal condicionado”

Contexto

Problema Inverso: hallar x en

$$Tx = y \quad (P)$$

donde $T : \mathcal{X} \rightarrow \mathcal{Y}$ tal que $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, \mathcal{X}, \mathcal{Y} espacios de Hilbert; $\mathcal{R}(T)$ no cerrado; y : dato.



El problema (P) es “mal condicionado” $\Rightarrow T^\dagger$ no acotado \Rightarrow **REGULARIZACIÓN**

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Método de Tikhonov-Phillips generalizado: $\min_{x \in \mathcal{D}(W)} \{ \|Tx - y\|^2 + \alpha W(x) \}$,

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Método de Tikhonov-Phillips generalizado: $\min_{x \in \mathcal{D}(W)} \{ \|Tx - y\|^2 + \alpha W(x) \}$,

- bondad de ajuste;

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Método de Tikhonov-Phillips generalizado: $\min_{x \in \mathcal{D}(W)} \{ \|Tx - y\|^2 + \alpha W(x) \}$,

- bondad de ajuste;
- **introducir estabilidad;**

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Método de Tikhonov-Phillips generalizado: $\min_{x \in \mathcal{D}(W)} \{ \|Tx - y\|^2 + \alpha W(x) \}$,

- bondad de ajuste;
- introducir estabilidad;
- incorporar información “a-priori” sobre la solución;

Contexto

Método de Tikhonov-Phillips de orden cero (clásico): $\min_{x \in \mathcal{X}} \{ \|Tx - y\|^2 + \alpha \|x\|^2 \}$,

Método de Tikhonov-Phillips generalizado: $\min_{x \in \mathcal{D}(W)} \{ \|Tx - y\|^2 + \alpha W(x) \}$,

- bondad de ajuste;
- introducir estabilidad;
- incorporar información “a-priori” sobre la solución;
- **balance entre bondad de ajuste y estabilidad.**

Enfoque estadístico

VARIABLES ALEATORIAS: $Y = AX + E$

Enfoque estadístico

VARIABLES ALEATORIAS: $Y = AX + E$

**Solución del Problema Inverso:
DISTRIBUCIÓN DE PROBABILIDAD**

Fórmula de Bayes

Teorema

Sea $X \in \mathbb{R}^n$ una variable aleatoria con **densidad a-priori** $\pi_{pr}(x)$. Sea y_{obs} una observación de la variable aleatoria $Y \in \mathbb{R}^k$ tal que $\pi(y_{obs}) > 0$. Entonces, la **distribución de probabilidad a-posteriori** de X , dado el dato y_{obs} es

$$\pi_{post}(x) = \pi(x|y_{obs}) = \frac{\pi(y_{obs}|x)\pi_{pr}(x)}{\pi(y_{obs})} \propto \pi(y_{obs}|x)\pi_{pr}(x).$$

Fórmula de Bayes

Teorema

Sea $X \in \mathbb{R}^n$ una variable aleatoria con *densidad a-priori* $\pi_{pr}(x)$. Sea y_{obs} una observación de la variable aleatoria $Y \in \mathbb{R}^k$ tal que $\pi(y_{obs}) > 0$. Entonces, la *distribución de probabilidad a-posteriori* de X , dado el dato y_{obs} es

$$\pi_{post}(x) = \pi(x|y_{obs}) = \frac{\pi(y_{obs}|x)\pi_{pr}(x)}{\pi(y_{obs})} \propto \pi(y_{obs}|x)\pi_{pr}(x).$$

1) Elegir una π_{pr} adecuada.

Fórmula de Bayes

Teorema

Sea $X \in \mathbb{R}^n$ una variable aleatoria con *densidad a-priori* $\pi_{pr}(x)$. Sea y_{obs} una observación de la variable aleatoria $Y \in \mathbb{R}^k$ tal que $\pi(y_{obs}) > 0$. Entonces, la *distribución de probabilidad a-posteriori* de X , dado el dato y_{obs} es

$$\pi_{post}(x) = \pi(x|y_{obs}) = \frac{\pi(y_{obs}|x)\pi_{pr}(x)}{\pi(y_{obs})} \propto \pi(y_{obs}|x)\pi_{pr}(x).$$

- 1) Elegir una π_{pr} adecuada.
- 2) Hallar $\pi(y_{obs}|x)$.

Fórmula de Bayes

Teorema

Sea $X \in \mathbb{R}^n$ una variable aleatoria con **densidad a-priori** $\pi_{pr}(x)$. Sea y_{obs} una observación de la variable aleatoria $Y \in \mathbb{R}^k$ tal que $\pi(y_{obs}) > 0$. Entonces, la **distribución de probabilidad a-posteriori** de X , dado el dato y_{obs} es

$$\pi_{post}(x) = \pi(x|y_{obs}) = \frac{\pi(y_{obs}|x)\pi_{pr}(x)}{\pi(y_{obs})} \propto \pi(y_{obs}|x)\pi_{pr}(x).$$

- 1) Elegir una π_{pr} adecuada.
- 2) Hallar $\pi(y_{obs}|x)$.
- 3) Explorar π_{post} .

Estimadores

Definición

Dada la densidad de probabilidad *a-posteriori* $\pi(x|y)$ de la variable aleatoria incógnita $X \in \mathbb{R}^n$, definimos:

1) El **estimador máximo a-posteriori** como

$$x_{\text{MAP}} \doteq \arg \max_{x \in \mathbb{R}^n} \pi(x|y).$$

Estimadores

Definición

Dada la densidad de probabilidad *a-posteriori* $\pi(x|y)$ de la variable aleatoria incógnita $X \in \mathbb{R}^n$, definimos:

- 1) El **estimador máximo a-posteriori** como

$$x_{\text{MAP}} \doteq \arg \max_{x \in \mathbb{R}^n} \pi(x|y).$$

- 2) El **estimador media condicional** como

$$x_{\text{CM}} \doteq E(x|y) = \int_{\mathbb{R}^n} x \pi(x|y) dx,$$

bajo el supuesto de que la integral exista.

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma} (Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma} (Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

$$x_{\text{MAP}} \doteq \underset{x \in \mathbb{R}^n}{\operatorname{argmax}} \pi(x|y)$$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

$$\begin{aligned} x_{\text{MAP}} &\doteq \operatorname{argmax}_{x \in \mathbb{R}^n} \pi(x|y) \\ &= \operatorname{argmax}_{x \in \mathbb{R}^n} \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right) \end{aligned}$$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

$$\begin{aligned}
 x_{\text{MAP}} &\doteq \operatorname{argmax}_{x \in \mathbb{R}^n} \pi(x|y) \\
 &= \operatorname{argmax}_{x \in \mathbb{R}^n} \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left(\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right)
 \end{aligned}$$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

$$\begin{aligned}
 x_{\text{MAP}} &\doteq \operatorname{argmax}_{x \in \mathbb{R}^n} \pi(x|y) \\
 &= \operatorname{argmax}_{x \in \mathbb{R}^n} \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left(\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left(\|Ax - y\|^2 + \alpha \|Rx\|^2 \right) \text{ donde } \alpha = \sigma^2
 \end{aligned}$$

Relación entre los enfoques clásico y estadístico

- Modelo: $Y = AX + E$, $X \in \mathbb{R}^n$ e $Y, E \in \mathbb{R}^k$, $A \in \mathbb{R}^{k \times n}$
- $X \sim \mathcal{N}(0, \Gamma)$
- $E \sim \mathcal{N}(0, \sigma^2 I_k)$
- $\pi(x|y) \propto \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right)$ donde $\Gamma^{-1} = R^T R$.

$$\begin{aligned}
 x_{\text{MAP}} &\doteq \operatorname{argmax}_{x \in \mathbb{R}^n} \pi(x|y) \\
 &= \operatorname{argmax}_{x \in \mathbb{R}^n} \exp\left(-\frac{1}{2} \left[\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right]\right) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left(\left\| \frac{1}{\sigma}(Ax - y) \right\|^2 + \|Rx\|^2 \right) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left(\|Ax - y\|^2 + \alpha \|Rx\|^2 \right) \text{ donde } \alpha = \sigma^2 \\
 &= x_{\text{TIKG}} \quad \text{▶ TG}
 \end{aligned}$$

Restauración de imágenes

$$y = Tx \quad (P)$$

Restauración de imágenes

$$y = Tx \quad (P)$$

$$g(s, t) = \int_0^1 \int_0^1 a(s, t, s', t') f(s', t') ds' dt',$$

Restauración de imágenes

$$y = Tx \quad (P)$$

$$g(s, t) = \int_0^1 \int_0^1 a(s, t, s', t') f(s', t') ds' dt',$$

a está dado por la Gaussiana bidimensional

$$a(s, t, s', t') = \frac{1}{2\pi\sigma_h\sigma_v} \exp\left(-\frac{1}{2}\left(\frac{s-s'}{\sigma_h}\right)^2 - \frac{1}{2}\left(\frac{t-t'}{\sigma_v}\right)^2\right),$$

Restauración de imágenes

$$y = Tx \quad (P)$$

$$g(s, t) = \int_0^1 \int_0^1 a(s, t, s', t') f(s', t') ds' dt',$$

a está dado por la Gaussiana bidimensional

$$a(s, t, s', t') = \frac{1}{2\pi\sigma_h\sigma_v} \exp\left(-\frac{1}{2}\left(\frac{s-s'}{\sigma_h}\right)^2 - \frac{1}{2}\left(\frac{t-t'}{\sigma_v}\right)^2\right),$$

$$Y = AX + E \text{ donde } X, Y, E \in \mathbb{R}^{n^2}, A \text{ de orden } n^2 \times n^2.$$

Detección de anomalías

(Kaipio, J. and Somersalo, E., *Estimating anomalies from indirect observations*, Journal of Computational Physics, 181:398-406, 2002.)

$$X = U + V$$

Detección de anomalías

(Kaipio, J. and Somersalo, E., *Estimating anomalies from indirect observations*, Journal of Computational Physics, 181:398-406, 2002.)

$$X = U + V$$



PARTE REGULAR

Detección de anomalías

(Kaipio, J. and Somersalo, E., *Estimating anomalies from indirect observations*, Journal of Computational Physics, 181:398-406, 2002.)

$$X = U + V$$



PARTE REGULAR

PARTE ANÓMALA

Detección de anomalías

(Kaipio, J. and Somersalo, E., *Estimating anomalies from indirect observations*, Journal of Computational Physics, 181:398-406, 2002.)

$$X = U + V$$



PARTE REGULAR

PARTE ANÓMALA

OBJETIVO: ESTIMAR U y V

Detección de anomalías

$$X = U + V$$

Detección de anomalías

$$X = U + V$$

$U \sim$ prior de suavidad

Detección de anomalías

$$X = U + V$$

$$U \sim \text{prior de suavidad} \rightsquigarrow U \sim \pi_1(u) \propto \exp\left(-\frac{\alpha}{2} \|Lu\|^2\right)$$

Detección de anomalías

$$X = U + V$$

$U \sim$ prior de suavidad $\rightsquigarrow U \sim \pi_1(u) \propto \exp\left(-\frac{\alpha}{2} \|Lu\|^2\right)$

$V \sim$ prior de impulso

Detección de anomalías

$$X = U + V$$

$$U \sim \text{prior de suavidad} \rightsquigarrow U \sim \pi_1(u) \propto \exp\left(-\frac{\alpha}{2} \|Lu\|^2\right)$$

$$V \sim \text{prior de impulso} \rightsquigarrow V \sim \pi_2(v) \propto \pi_+(v) \exp(-\gamma \|v\|_1)$$

Detección de anomalías

$$\begin{aligned} Y &= AX + E \\ &= AU + AV + E, \end{aligned}$$

Detección de anomalías

$$\begin{aligned} Y &= AX + E \\ &= AU + AV + E, \quad E \sim \mathcal{N}(0, \sigma_e^2 I_{n^2} \doteq \Gamma_e) \end{aligned}$$

Detección de anomalías

$$\begin{aligned} Y &= AX + E \\ &= AU + AV + E, \quad E \sim \mathcal{N}(0, \sigma_e^2 I_{n^2} \doteq \Gamma_e) \end{aligned}$$



$$\pi(u, y, v)$$

Detección de anomalías

$$\pi(u, y, v) = \pi(u, y|v)\pi_2(v)$$

Detección de anomalías

$$\pi(u, y, v) = \pi(u, y|v)\pi_2(v)$$

$$\propto \exp\left(-\frac{1}{2}[u^T(y - Av)^T]\Gamma^{-1}\begin{bmatrix} u \\ y - Av \end{bmatrix}\right) \pi_+(v) \exp(-\gamma \|v\|_1)$$

donde $\Gamma = \begin{pmatrix} \Gamma_u & \Gamma_u A^T \\ A\Gamma_u & A\Gamma_u A^T + \Gamma_e \end{pmatrix}$,

Detección de anomalías

$$\pi(u, y, v) = \pi(u, y|v)\pi_2(v)$$

$$\propto \exp\left(-\frac{1}{2}[u^T(y - Av)^T]\Gamma^{-1}\begin{bmatrix} u \\ y - Av \end{bmatrix}\right) \pi_+(v) \exp(-\gamma \|v\|_1)$$



Densidad Condicional
 $\pi(u|v, y)$

Densidad Condicional
 $\pi(v|u, y)$

Densidades Condicionales

(Kaipio, J. and Somersalo, E., *Statistical and computational inverse problems*, volume 160 of Applied Mathematical Sciences. Springer-Verlag, New York, 2005.)

$$\pi(u|v, y) \propto \exp\left(-\frac{1}{2}(u - u_0)^T \tilde{\Gamma}_{22}^{-1}(u - u_0)\right),$$

- $\Gamma = \begin{pmatrix} \Gamma_u & \Gamma_u A^T \\ A\Gamma_u & A\Gamma_u A^T + \Gamma_e \end{pmatrix} \doteq \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix},$
- $u_0 = \Gamma_{12}\Gamma_{22}^{-1}(y - Av),$
- $\tilde{\Gamma}_{22} \doteq \Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}$ es el complemento de Schur de $\Gamma_{22}.$

Densidades Condicionales

$$\begin{aligned}\pi(v|u, y) &\propto \pi_+(v) \exp\left(-\frac{1}{2}(v^T A^T R_{22} A v - 2v^T (A^T R_{21} u + A^T R_{22} y)) - \gamma \|v\|_1\right) \\ &= \pi_+(v) \exp\left(-\frac{1}{2}(v^T B v - 2v^T q)\right),\end{aligned}$$

- R_{ij} denotan los bloques de la matriz $\Gamma^{-1} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$,
- $B \doteq A^T R_{22} A$,
- $q \doteq A^T R_{21} u + A^T R_{22} y - \gamma \mathbf{1}$, $\mathbf{1} = [1, \dots, 1]^T$, pues $v_k \geq 0$ implica que $\|v\|_1 = \mathbf{1}^T v$.

Densidades Condicionales

$$\pi(u|v, y) \sim \mathcal{N}(u_0, \tilde{\Gamma}_{22})$$

$$\pi(v|u, y) \propto \mathcal{N}(-, -)$$

Densidades Condicionales

$$\pi(u|v, y) \sim \mathcal{N}(u_0, \tilde{\Gamma}_{22})$$

$$\pi(v|u, y) \propto \mathcal{N}(-, -)$$

Estimador U_{CM}

Estimador V_{CM}

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx)$$

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx) \approx \sum_{j=1}^N w_j f(x_j) \quad (\text{Métodos de cuadratura})$$

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx) \approx \sum_{j=1}^N w_j f(x_j) \quad (\text{Métodos de cuadratura})$$

$X \in \mathbb{R}^n$ una variable aleatoria con distribución de probabilidad μ

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx) \approx \sum_{j=1}^N w_j f(x_j) \quad (\text{Métodos de cuadratura})$$

$X \in \mathbb{R}^n$ una variable aleatoria con distribución de probabilidad μ

$$\{x_1, x_2, \dots, x_N\} \subset \mathbb{R}^n$$

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx) \approx \sum_{j=1}^N w_j f(x_j) \quad (\text{Métodos de cuadratura})$$

$X \in \mathbb{R}^n$ una variable aleatoria con distribución de probabilidad μ

$$\{x_1, x_2, \dots, x_N\} \subset \mathbb{R}^n$$

$$\int_{\mathbb{R}^n} f(x)\mu(dx) = E[f(X)] \approx \frac{1}{N} \sum_{j=1}^N f(x_j) \quad (\text{Métodos de Monte Carlo})$$

μ : medida de probabilidad en \mathbb{R}^n , $f \in L^1(\mu(dx))$.

$$\int_{\mathbb{R}^n} f(x)\mu(dx) \approx \sum_{j=1}^N w_j f(x_j) \quad (\text{Métodos de cuadratura})$$

$X \in \mathbb{R}^n$ una variable aleatoria con distribución de probabilidad μ

$$\{x_1, x_2, \dots, x_N\} \subset \mathbb{R}^n$$

$$\int_{\mathbb{R}^n} f(x)\mu(dx) = E[f(X)] \approx \frac{1}{N} \sum_{j=1}^N f(x_j) \quad (\text{Métodos de Monte Carlo})$$



Métodos de Monte Carlo con cadenas de Markov

Métodos de Monte Carlo con cadenas de Markov

Un mapeo $P : \mathbb{R}^n \times \mathfrak{B} \rightarrow [0, 1]$ se llama *núcleo de probabilidad de transición* si

Métodos de Monte Carlo con cadenas de Markov

Un mapeo $P : \mathbb{R}^n \times \mathfrak{B} \rightarrow [0, 1]$ se llama *núcleo de probabilidad de transición* si

- para cada $B \in \mathfrak{B}$, el mapeo $\mathbb{R}^n \rightarrow [0, 1], x \rightarrow P(x, B)$ es una función medible;

Métodos de Monte Carlo con cadenas de Markov

Un mapeo $P : \mathbb{R}^n \times \mathfrak{B} \rightarrow [0, 1]$ se llama *núcleo de probabilidad de transición* si

- 1 para cada $B \in \mathfrak{B}$, el mapeo $\mathbb{R}^n \rightarrow [0, 1], x \rightarrow P(x, B)$ es una función medible;
- 2 para cada $x \in \mathbb{R}^n$, el mapeo $\mathfrak{B} \rightarrow [0, 1], B \rightarrow P(x, B)$ es una distribución de probabilidad.

Métodos de Monte Carlo con cadenas de Markov

Un mapeo $P : \mathbb{R}^n \times \mathfrak{B} \rightarrow [0, 1]$ se llama *núcleo de probabilidad de transición* si

- 1 para cada $B \in \mathfrak{B}$, el mapeo $\mathbb{R}^n \rightarrow [0, 1], x \rightarrow P(x, B)$ es una función medible;
- 2 para cada $x \in \mathbb{R}^n$, el mapeo $\mathfrak{B} \rightarrow [0, 1], B \rightarrow P(x, B)$ es una distribución de probabilidad.

Una *cadena de Markov de tiempo homogéneo* con núcleo de transición P es un proceso estocástico $\{X_j\}_{j=1}^{\infty}$, $X_j \in \mathbb{R}^n$, que satisface

$$\mu_{x_{j+1}}(B_{j+1} | x_1, \dots, x_j) = \mu_{x_{j+1}}(B_{j+1} | x_j) = P(x_j, B_{j+1}).$$

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico*

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico* y μ es una *medida invariante* del núcleo P

$$\text{es decir } \mu P(B) = \int_{\mathbb{R}^n} P(x, B) \mu(dx) = \mu(B)$$

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico* y μ es una *medida invariante* del núcleo P . Entonces para todo $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} P^{(N)}(x, B) = \mu(B) \quad \text{para todo } B \in \mathfrak{B}.$$

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico* y μ es una *medida invariante* del núcleo P . Entonces para todo $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} P^{(N)}(x, B) = \mu(B) \quad \text{para todo } B \in \mathfrak{B}.$$

¿Cómo explorar una distribución de probabilidad μ dada?

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico* y μ es una *medida invariante* del núcleo P . Entonces para todo $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} P^{(N)}(x, B) = \mu(B) \quad \text{para todo } B \in \mathfrak{B}.$$

¿Cómo explorar una distribución de probabilidad μ dada?

Construir P irreducible, aperiódico tal que μ sea invariante para P , para obtener $\{X_1, X_2, \dots, X_N\}$.

Métodos de Monte Carlo con cadenas de Markov

Proposición: Sea μ una medida de probabilidad en \mathbb{R}^n y $\{X_j\}$ una cadena de Markov de tiempo homogéneo con núcleo P . Supongamos además que P es *irreducible* y *aperiódico* y μ es una *medida invariante* del núcleo P . Entonces para todo $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} P^{(N)}(x, B) = \mu(B) \quad \text{para todo } B \in \mathfrak{B}.$$

¿Cómo explorar una distribución de probabilidad μ dada?

Construir P irreducible, aperiódico tal que μ sea invariante para P , para obtener $\{x_1, x_2, \dots, x_N\}$.



Metropolis-Hastings (Metropolis, N. *et al*, 1953; Hastings, W. K., 1970)

Muestreador de Gibbs (Geman, D. and Geman, S., 1984)

Muestreador de Gibbs

Algoritmo de muestreo:

Inicializar $(u, v) = (u^0, v^0)$

setear $k = 0$

Hasta llegar al tamaño de muestra deseada:

$(u, v) = (u^k, v^k)$

muestrear u^{k+1} de $\pi(u|v, y)$

para $p = 1 : n^2$

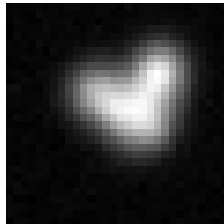
muestrear $v_p^{k+1} \geq 0$ de $\pi(v_p|u, v_{-p}, y)$

end

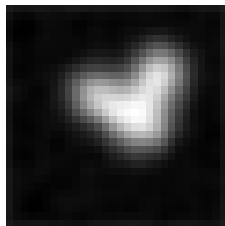
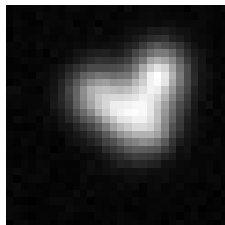
$k \leftarrow k + 1$

end

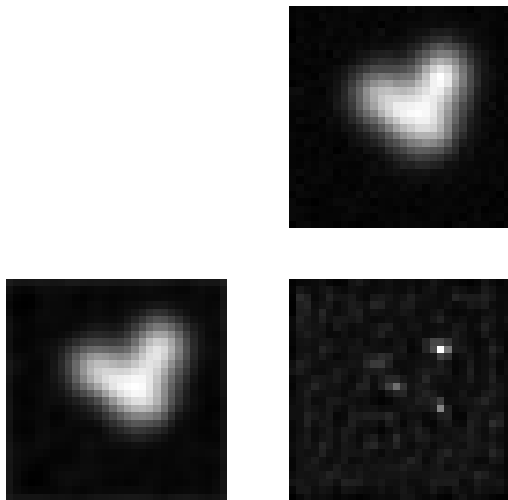
Resultados numéricos



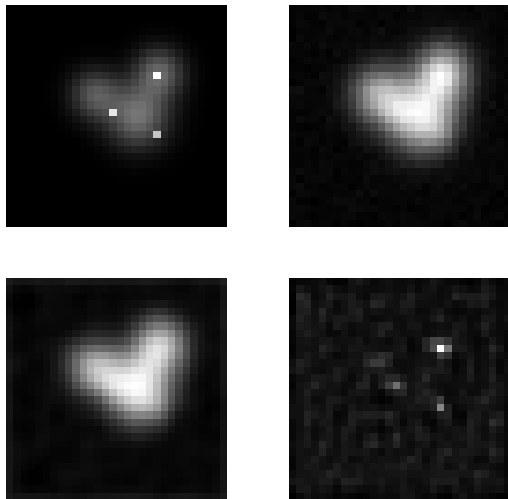
Resultados numéricos



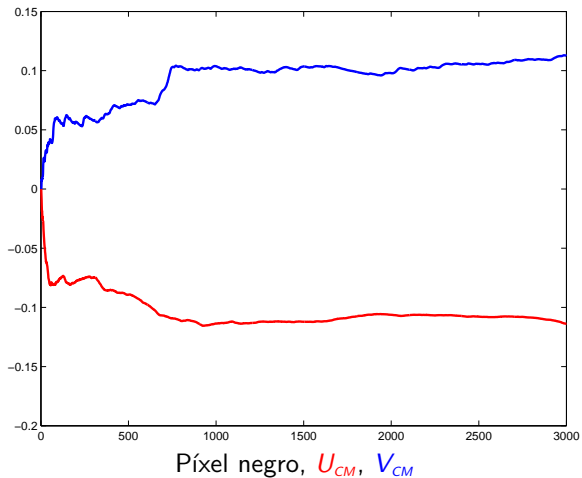
Resultados numéricos



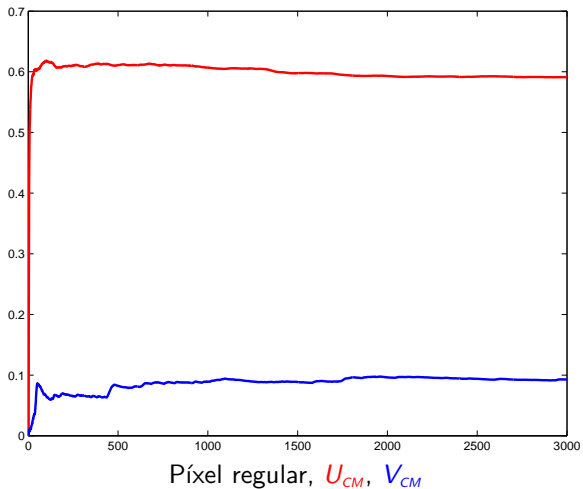
Resultados numéricos



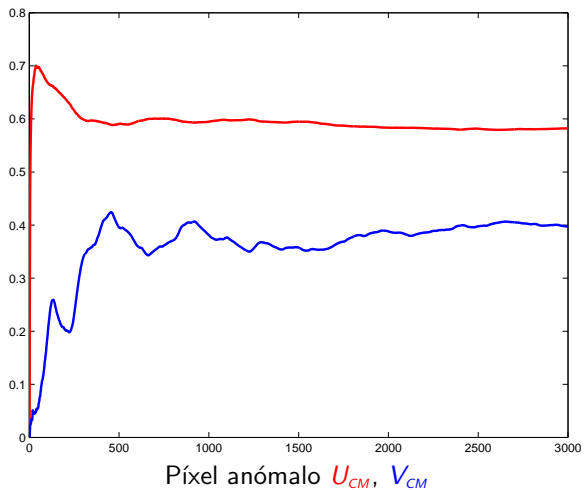
Resultados numéricos: período de ablande



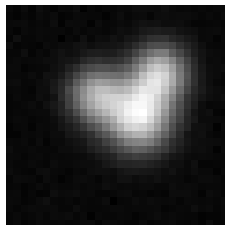
Resultados numéricos: período de ablande



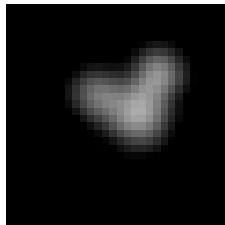
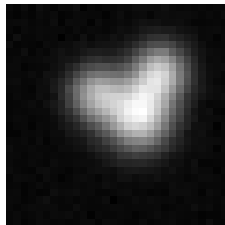
Resultados numéricos: período de ablande



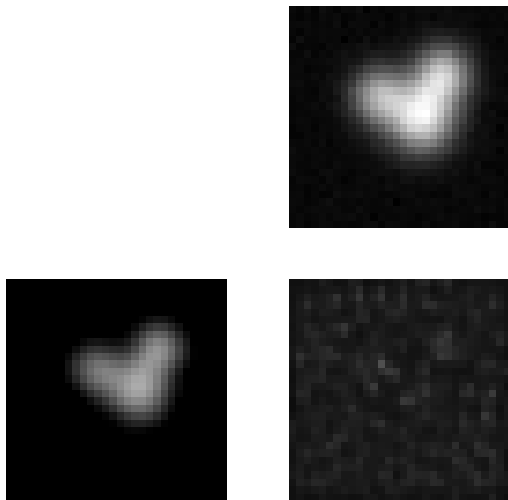
Resultados numéricos ($X = U$)

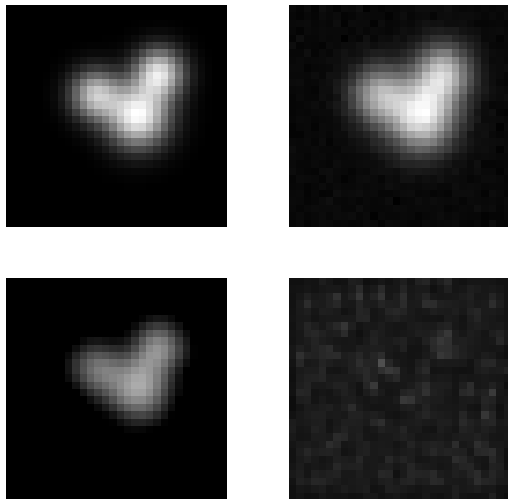


Resultados numéricos ($X = U$)






Resultados numéricos ($X = U$)



Resultados numéricos ($X = U$)

Bibliografía

-  CALVETTI, D. AND SOMERSALO, E., *Introduction to Bayesian Scientific Computing. Ten Lectures on Subjective Computing*, volume 2 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2007.
-  ENGL, H. W.; HANKE, M. AND NEUBAUER, A., *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
-  KAIPIO, J. AND SOMERSALO, E., *Estimating anomalies from indirect observations*, *Journal of Computational Physics*, 181:398-406, 2002.
-  KAIPIO, J. AND SOMERSALO, E., *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.

¡Gracias por su atención!