Processo di Dirichlet

Luigi Amedeo Bianchi

SNS, Pisa

Parma, 22 dicembre 2010

Dove siamo?

Definizione

Dati $(\mathcal{X}, \mathcal{A})$ ed α,

Definizione

Dati $(\mathcal{X}, \mathcal{A})$ ed α, μ è un *processo di Dirichlet* di parametro α su $\mathcal{X}$ se è un processo stocastico $(\mu_A)_{A\in\mathcal{A}}$ a valori in $[0, +\infty]$ tale che

$$\mu_{B_1,\ldots,B_k} \sim \mathrm{Dir}\left(\alpha\left(B_1\right), \ldots, \alpha\left(B_k\right)\right),$$

Definizione

Dati $(\mathcal{X}, \mathcal{A})$ ed α, μ è un *processo di Dirichlet* di parametro α su $\mathcal{X}$ se è un processo stocastico $(\mu_A)_{A \in \mathcal{A}}$ a valori in $[0, +\infty]$ tale che

$$\mu_{B_1,\dots,B_k} \sim \mathrm{Dir}\left(\alpha\left(B_1\right), \dots, \alpha\left(B_k\right)\right),$$

cioè ha densità

$$f(x_1, \dots, x_{k-1}) = \frac{\Gamma\left(\sum_{i=1}^{k+1} \alpha_i\right)}{\prod_{i=1}^{k+1} \Gamma\left(\alpha_i\right)} x_1^{\alpha_1 - 1} \cdots x_{n-1}^{\alpha_{k-1} - 1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{\alpha_k - 1} \mathbf{1}_\Delta$$

in cui $\alpha_i = \alpha(B_i)$.

Kolmogorov

Definizione

Una famiglia $\{\mathbb{P}_{t_1,\dots,t_k}, k \in \mathbb{N}, t_1, \dots, t_k \in T\}$ di misure di probabilità su $\mathcal{X}^k$ si dice *consistente* se

$$\mathbb{P}_{t_1,\dots,t_k}\Big|_{\{t_{j_1},\dots,t_{j_h}\}} = \mathbb{P}_{t_{j_1},\dots,t_{j_h}}.$$

Kolmogorov

Definizione

Una famiglia $\{\mathbb{P}_{t_1,\ldots,t_k}, k \in \mathbb{N}, t_1, \ldots, t_k \in T\}$ di misure di probabilità su $\mathcal{X}^k$ si dice *consistente* se

$$\mathbb{P}_{t_1,\ldots,t_k}\Big|_{\{t_{j_1},\ldots,t_{j_h}\}} = \mathbb{P}_{t_{j_1},\ldots,t_{j_h}}.$$

Teorema (Kolmogorov)

Esistono uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ *ed un processo stocastico* $(X_t)_{t\in T}$ *su* Ω *a valori in* $\mathcal{X}$ *tali che le* $\mathbb{P}_{t_1,\ldots,t_k}$ *siano le distribuzioni finito dimensionali di* $\{X_t\}_{t\in T}$*, ossia*

$$\mathbb{P}_{t_1,\ldots,t_k} = \mathbb{P}\left(X_{t_1}, \ldots, X_{t_k}\right)^{-1}.$$

Distribuzioni finito-dimensionali

Dalla definizione data per il Processo di Dirichlet abbiamo le distribuzioni finito-dimensionali, ma. . .

Distribuzioni finito-dimensionali

Dalla definizione data per il Processo di Dirichlet abbiamo le distribuzioni finito-dimensionali, ma. . .

- sono assegnate solo sulle partizioni
- non sappiamo se siano compatibili.

Distribuzioni finito-dimensionali

Dalla definizione data per il Processo di Dirichlet abbiamo le distribuzioni finito-dimensionali, ma. . .

- sono assegnate solo sulle partizioni
- non sappiamo se siano compatibili.

In realtà:

- abbiamo un metodo per passare dalle partizioni alle famiglie finite

Distribuzioni finito-dimensionali

Dalla definizione data per il Processo di Dirichlet abbiamo le distribuzioni finito-dimensionali, ma. . .

- sono assegnate solo sulle partizioni
- non sappiamo se siano compatibili.

In realtà:

- abbiamo un metodo per passare dalle partizioni alle famiglie finite
- ci basta provare la compatibilità sulle partizioni.

Distribuzioni finito-dimensionali

Dalla definizione data per il Processo di Dirichlet abbiamo le distribuzioni finito-dimensionali, ma. . .

- sono assegnate solo sulle partizioni
- non sappiamo se siano compatibili.

In realtà:

- abbiamo un metodo per passare dalle partizioni alle famiglie finite
- ci basta provare la compatibilità sulle partizioni.

La definizione è quindi ben data.

Un processo di Dirichlet ha $\mathcal{A}$ come insieme dei parametri, introduciamo la notazione $\mu(A) = \mu_A$.

Proposizione

Siano μ *un processo di Dirichlet su* $(\mathcal{X}, \mathcal{A})$ *di parametro* α *e* $A \in \mathcal{A}$. *Se* $\alpha(A) = 0$ *allora* $\mu(A) = 0$ *q.c.; se* $\alpha(A) > 0$ *allora* $\mu(A) > 0$ *q.c. Inoltre* $\mathbb{E}[\mu(A)] = \frac{\alpha(A)}{\alpha(\mathcal{X})}$, *in cui la speranza è rispetto alla misura di probabilità* $\mathbb{P}$.

Criterio di Kallenberg

Teorema

Sia $\{\mathbb{P}_{B_1,\ldots,B_k}\}$, *una famiglia consistente di misure di probabilità. Allora esiste una misura aleatoria* μ *sullo spazio* $\mathcal{X}$ *che soddisfa*

$$\mathbb{P}\left(\mu\left(B_1\right),\ldots,\mu\left(B_k\right)\right)^{-1}=\mathbb{P}_{B_1,\ldots,B_k}$$

se e solo se

(i) $\mathbb{P}_{A,B,A\cup B}\left((x,y,z):\ x+y=z\right)=1$ *per tutti gli* A, B *disgiunti tra loro,*

(ii) $\mathbb{P}_{B_n}\rightharpoonup\delta_0$ *per ogni* $\left(B_n\right)_n$ *con* $B_n\downarrow\emptyset$.

Misura aleatoria

Osservazione

La (i) è una conseguenza del metodo usato per costruire le distribuzioni finito dimensionali.

Misura aleatoria

Osservazione

La (i) è una conseguenza del metodo usato per costruire le distribuzioni finito dimensionali.

La (ii) segue da:

Misura aleatoria

Osservazione

La (i) è una conseguenza del metodo usato per costruire le distribuzioni finito dimensionali.

La (ii) segue da:

Proposizione

Sia μ *un processo di Dirichlet su* $(\mathcal{X}, \mathcal{A})$ *di parametro* α*. Allora per una fissata successione decrescente di insiemi misurabili* $A_n \downarrow \emptyset$ *abbiamo* $\mu(A_n) \to 0$ *quasi certamente.*

Misura aleatoria

Osservazione

La (i) è una conseguenza del metodo usato per costruire le distribuzioni finito dimensionali.

La (ii) segue da:

Proposizione

Sia μ *un processo di Dirichlet su* $(\mathcal{X}, \mathcal{A})$ *di parametro* α. *Allora per una fissata successione decrescente di insiemi misurabili* $A_n \downarrow \emptyset$ *abbiamo* $\mu(A_n) \to 0$ *quasi certamente.*

Allora μ è misura aleatoria di probabilità.

Misura aleatoria

Osservazione

La (i) è una conseguenza del metodo usato per costruire le distribuzioni finito dimensionali.

La (ii) segue da:

Proposizione

Sia μ un processo di Dirichlet su $(\mathcal{X}, \mathcal{A})$ di parametro α. Allora per una fissata successione decrescente di insiemi misurabili $A_n \downarrow \emptyset$ abbiamo $\mu(A_n) \to 0$ quasi certamente.

Allora μ è misura aleatoria di probabilità.

$$\mu_A(\omega) \qquad \text{vs.} \qquad \mu_\omega(A).$$

Distribuzioni discrete

Teorema

Sia μ *un processo di Dirichlet di parametro* α *su* $(\mathcal{X}, \mathcal{A})$ *e sia* $(X_1, \ldots, X_n)$ *un campione di lunghezza n estratto da* μ*. Allora*

$$\mu | X_1, \ldots, X_n \sim \mathrm{DP}\left(\alpha + \sum_{i=1}^{n} \delta_{X_i}\right).$$

Distribuzioni discrete

Teorema

Sia μ un processo di Dirichlet di parametro α su $(\mathcal{X}, \mathcal{A})$ e sia $(X_1, \ldots, X_n)$ un campione di lunghezza n estratto da μ. Allora

$$\mu | X_1, \ldots, X_n \sim \mathrm{DP}\left(\alpha + \sum_{i=1}^{n} \delta_{X_i}\right).$$

Da questa proprietà le costruzioni alternative

- via urne di Pólya

Distribuzioni discrete

Teorema

Sia μ un processo di Dirichlet di parametro α su $(\mathcal{X}, \mathcal{A})$ e sia $(X_1, \ldots, X_n)$ un campione di lunghezza n estratto da μ. Allora

$$\mu | X_1, \ldots, X_n \sim \mathrm{DP}\left(\alpha + \sum_{i=1}^{n} \delta_{X_i}\right).$$

Da questa proprietà le costruzioni alternative

- via urne di Pólya
- via stick-breaking.

Dove siamo?

Entropia relativa

Consideriamo lo spazio $(\mathcal{P}(M), d_W)$ e, su di esso, l'entropia relativa

Entropia relativa

Consideriamo lo spazio $(\mathcal{P}(M), d_W)$ e, su di esso, l'entropia relativa

$$\mathrm{Ent}(\mu) = \begin{cases} \int_M \rho \log \rho \, dx & \text{se } d\mu(x) \ll dx \quad \rho(x) = \frac{d\mu(x)}{dx} \\ +\infty & \text{altrimenti.} \end{cases}$$

Entropia relativa

Consideriamo lo spazio $(\mathcal{P}(M), d_W)$ e, su di esso, l'entropia relativa

$$\mathrm{Ent}(\mu) = \begin{cases} \int_M \rho \log \rho \, dx & \text{se } d\mu(x) \ll dx \quad \rho(x) = \frac{d\mu(x)}{dx} \\ +\infty & \text{altrimenti.} \end{cases}$$

Misura entropica

Caratterizzazione formale:

$$d\mathbb{P}^{\beta}(\mu) = \frac{1}{Z} e^{-\beta \cdot \mathrm{Ent}(\mu)} d\mathbb{P}^{0}(\mu).$$

Perché?

Perché?

- flusso gradiente in $(\mathcal{P}(M), d_W)$ rispetto al funzionale d'entropia

Perché?

- flusso gradiente in $(\mathcal{P}(M), d_W)$ rispetto al funzionale d'entropia per l'equazione del calore su M;

Perché?

- flusso gradiente in $(\mathcal{P}(M), d_W)$ rispetto al funzionale d'entropia per l'equazione del calore su M;
- flusso di altri funzionali per PDE dissipative su M;

Perché?

- flusso gradiente in $(\mathcal{P}(M), d_W)$ rispetto al funzionale d'entropia per l'equazione del calore su M;
- flusso di altri funzionali per PDE dissipative su M;
- variante stocastica del moto gradiente legata a $\mathbb{P}^\beta$ su $\mathcal{P}(M)$.

Costruzione per $M = [0, 1]$

Sia $\mathcal{G}_0$ lo spazio delle funzioni non decrescenti e continue a destra $g : [0, 1[\to [0, 1]$;

Costruzione per $M = [0, 1]$

Sia $\mathcal{G}_0$ lo spazio delle funzioni non decrescenti e continue a destra $g : [0, 1[\to [0, 1]$;

Le mappe

- $\chi : \mathcal{G}_0 \to \mathcal{P}_0 \quad g \mapsto g_\# \mathrm{Leb}$

Costruzione per $M = [0, 1]$

Sia $\mathcal{G}_0$ lo spazio delle funzioni non decrescenti e continue a destra $g : [0, 1[\to [0, 1]$;

Le mappe

- $\chi : \mathcal{G}_0 \to \mathcal{P}_0 \quad g \mapsto g_{\#}\mathrm{Leb}$
- $\chi^{-1} : \mathcal{P}_0 \to \mathcal{G}_0 \quad \mu \mapsto g_\mu$, con

$$g_\mu(t) = \inf \{ s \in [0, 1] : \ \mu([0, s]) > t \} .$$

Costruzione per $M = [0,1]$

Sia $\mathcal{G}_0$ lo spazio delle funzioni non decrescenti e continue a destra $g : [0,1[\to [0,1]$;

Le mappe

- $\chi : \mathcal{G}_0 \to \mathcal{P}_0 \quad g \mapsto g_\# \mathrm{Leb}$
- $\chi^{-1} : \mathcal{P}_0 \to \mathcal{G}_0 \quad \mu \mapsto g_\mu$, con

$$g_\mu(t) = \inf\left\{s \in [0,1] : \ \mu([0,s]) > t\right\}.$$

- Per ogni $g \in \mathcal{G}_0$ possiamo definire *l'inversa destra* $g^{(-1)} \in \mathcal{G}_0$ come

$$g^{(-1)}(t) = \inf\left\{s \geq 0 : \ g(s) > t\right\}.$$

Euristica

Con le χ spostiamo il problema della misura da $\mathcal{P}_0$ a $\mathcal{G}_0$:

$$Q_0^\beta\left(dg\right) = \frac{1}{Z} e^{-\beta \cdot S(g)} \mathbb{Q}^0\left(dg\right).$$

Euristica

Con le χ spostiamo il problema della misura da $\mathcal{P}_0$ a $\mathcal{G}_0$:

$$Q_0^{\beta}(dg) = \frac{1}{Z} e^{-\beta \cdot S(g)} \mathbb{Q}^0 (dg) .$$

Euristica di Feynman

Ricordiamo la caratterizzazione euristica della misura di Wiener:

$$\mathbb{P}^W (dg) = \frac{1}{Z} e^{-\mathrm{H}(g)} \mathbb{P}^{\mathcal{G}^*} (dg) ,$$

Euristica

Con le χ spostiamo il problema della misura da $\mathcal{P}_0$ a $\mathcal{G}_0$:

$$Q_0^{\beta}(dg) = \frac{1}{Z} e^{-\beta \cdot S(g)} \mathbb{Q}^0 (dg) .$$

Euristica di Feynman

Ricordiamo la caratterizzazione euristica della misura di Wiener:

$$\mathbb{P}^W (dg) = \frac{1}{Z} e^{-\mathrm{H}(g)} \mathbb{P}^{\mathcal{G}^*} (dg) ,$$

per interpolazione abbiamo le leggi finito-dimensionali

$$\mathbb{P}^W (g_{t_1}, \ldots, g_{t_N}) = \frac{1}{Z_N} \exp\left(-\frac{1}{2} \sum_{i=1}^{N} \frac{|x_i - x_{i-1}|^2}{t_i - t_{i-1}} \right) p_N (\cdots) .$$

Interpoliamo l'entropia:

$$\begin{aligned} S_N(g) &= \inf\left\{S(\tilde{g}) : \ \tilde{g} \in \mathcal{G}^*, \ \tilde{g}(t_i) = g(t_i) \ \ \forall i\right\} \\ &= -\sum_{i=1}^{N+1} \log \frac{\left|g(t_i) - g(t_{i-1})\right|^2}{2(t_i - t_{i-1})} \cdot (t_i - t_{i-1}). \end{aligned}$$

Interpoliamo l'entropia:

$$S_N(g) = \inf\{S(\tilde{g}) : \ \tilde{g} \in \mathcal{G}^*, \ \tilde{g}(t_i) = g(t_i) \ \ \forall i\}$$
$$= -\sum_{i=1}^{N+1} \log \frac{|g(t_i) - g(t_{i-1})|^2}{2(t_i - t_{i-1})} \cdot (t_i - t_{i-1}).$$

Le leggi finito-dimensionali sono allora:

$$\mathbb{Q}_0^\beta(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N) =$$
$$= \frac{1}{Z_N} \exp\left(\beta \sum_{i=1}^{N+1} \log \frac{|x_i - x_{i-1}|^2}{t_i - t_{i-1}} \cdot (t_i - t_{i-1})\right) q_N(dx_1, \ldots, dx_N).$$

Determiniamo q_N

- proprietà di invarianza;

Determiniamo q_N

- proprietà di invarianza;
- sono misure, ma non di probabilità.

Determiniamo q_N

- proprietà di invarianza;
- sono misure, ma non di probabilità.

Lemma

Una famiglia di misure q_N, al variare di $N \in \mathbb{N}$, con densità continue soddisfa la proprietà di invarianza se e solo se

$$q_N(dx_1, \ldots, dx_N) = C^N \frac{dx_1 \cdots dx_N}{x_1 \cdot (x_2 - x_1) \cdots (x_N - x_{N-1}) \cdot (1 - x_N)},$$

per qualche costante $C \in \mathbb{R}_+$.

Allora

$$\mathbb{Q}_0^{\beta}\left(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N\right) =$$
$$= \frac{1}{Z_N} \prod_{i=1}^{N+1} (x_i - x_{i-1})^{\beta(t_i - t_{i-1}) - 1} \, dx_1 \cdots dx_N,$$

Allora

$$\mathbb{Q}_0^{\beta}\left(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N\right) =$$
$$= \frac{1}{Z_N} \prod_{i=1}^{N+1} \left(x_i - x_{i-1}\right)^{\beta(t_i - t_{i-1}) - 1} dx_1 \cdots dx_N,$$

che con

$$Z_N = \frac{\Gamma\left(\beta\right)}{\prod_{i=1}^{N+1} \Gamma\left(\beta\left(t_i - t_{i-1}\right)\right)}$$

è una famiglia di distribuzioni di Dirichlet sulla partizione $\{t_0 = 0, t_1, \ldots, t_N, t_{N+1} = 1\}$ dell'intervallo $[0, 1]$.

Allora

$$\mathbb{Q}_0^\beta\left(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N\right) = \\ = \frac{1}{Z_N} \prod_{i=1}^{N+1} (x_i - x_{i-1})^{\beta(t_i - t_{i-1}) - 1}\, dx_1 \cdots dx_N,$$

Dunque $\mathbb{Q}_0^\beta$ è proprio il processo di Dirichlet.

Definizione

Definizione

La *misura entropica* $\mathbb{P}_0^\beta$ su $\mathcal{P}_0 = \mathcal{P}([0,1])$ è definita come la misura immagine del processo di Dirichlet $\mathbb{Q}_0^\beta$ di parametro $\alpha = \beta \cdot \mathrm{Leb}$ su $\mathcal{G}_0$ rispetto alla mappa χ. Ciò significa che, per tutte le funzioni misurabili limitate $u : \mathcal{P}_0 \to \mathbb{R}$,

$$\int_{\mathcal{P}_0} u(\mu)\, d\mathbb{P}_0^\beta(\mu) = \int_{\mathcal{G}_0} u(g_\# \mathrm{Leb})\, d\mathbb{Q}_0^\beta(g).$$

Considerazioni

I passi

$$\mu \overset{f(x)=\mu([0,x])}{\longleftrightarrow} f \overset{g=f^{(-1)}}{\longleftrightarrow} g$$

Considerazioni

I passi

$$\mu \overset{f(x)=\mu([0,x])}{\longleftrightarrow} f \overset{g=f^{(-1)}}{\longleftrightarrow} g \overset{g(x)=\nu([0,x])}{\longleftrightarrow} \nu,$$

Considerazioni

I passi

$$\mu \overset{f(x)=\mu([0,x])}{\longleftrightarrow} f \overset{g=f^{(-1)}}{\longleftrightarrow} g \overset{g(x)=\nu([0,x])}{\longleftrightarrow} \nu,$$

ma anche, evidenziando le χ,

$$\mu \overset{g=\chi^{-1}(\mu)}{\longleftrightarrow} g \overset{g=f^{(-1)}}{\longleftrightarrow} f \overset{\nu=\chi(f)}{\longleftrightarrow} \nu.$$

Considerazioni

I passi

$$\mu \overset{f(x)=\mu([0,x])}{\longleftrightarrow} f \overset{g=f^{(-1)}}{\longleftrightarrow} g \overset{g(x)=\nu([0,x])}{\longleftrightarrow} \nu,$$

ma anche, evidenziando le χ,

$$\mu \overset{g=\chi^{-1}(\mu)}{\longleftrightarrow} g \overset{g=f^{(-1)}}{\longleftrightarrow} f \overset{\nu=\chi(f)}{\longleftrightarrow} \nu.$$

Generalizzando il passo $\mu = g_{\#}m$ con la mappa di Brenier, possiamo definire una *mappa di coniugio*

$$\mathfrak{C} : \mathcal{P}(M) \to \mathcal{P}(M), \quad \mu \mapsto \nu.$$

Definizione

Dati una varietà riemanniana compatta M ed un generico parametro $\beta > 0$, la *misura entropica*

$$\mathbb{P}^\beta = \mathfrak{C}_\# \mathbb{Q}^\beta$$

è la misura immagine del processo di Dirichlet $\mathbb{Q}^\beta$ su $\mathcal{P}(M)$ di parametro $\beta \cdot m$ rispetto alla mappa di coniugio $\mathfrak{C}$.

Definizione

Dati una varietà riemanniana compatta M ed un generico parametro $\beta > 0$, la *misura entropica*

$$\mathbb{P}^\beta = \mathfrak{C}_\# \mathbb{Q}^\beta$$

è la misura immagine del processo di Dirichlet $\mathbb{Q}^\beta$ su $\mathcal{P}(M)$ di parametro $\beta \cdot m$ rispetto alla mappa di coniugio $\mathfrak{C}$.

Notiamo che sia $\mathbb{P}^\beta$ che $\mathbb{Q}^\beta$ sono misure di probabilità sullo spazio compatto $\mathcal{P} = \mathcal{P}(M)$ delle misure di probabilità su M.

La strada fatta

One more thing...

Collaborazione

- Stesura di note o appunti;

- Sviluppo di codice;

- Scrittura di articoli.

Collaborazione

- Stesura di note o appunti;
- Sviluppo di codice;
- Scrittura di articoli.

Ma anche

Tracciamento

- Stesura di note o appunti;

- Sviluppo di codice;

- Scrittura di articoli.

Collaborazione

- Stesura di note o appunti;
- Sviluppo di codice;
- Scrittura di articoli.

Tracciamento

- Stesura di note o appunti;
- Sviluppo di codice;
- Scrittura di articoli.

Versionamento

Versionamento

Git: sistema di versionamento

Versionamento

Git: sistema di versionamento *distribuito*

Versionamento

Git: sistema di versionamento *distribuito*

http://git-scm.com/

Versionamento

Gitorius: web host per progetti collaborativi open source;

Versionamento

Gitorius:
in particolare una piattaforma per la condivisione di appunti scientifici si trova su

http://git.phc.unipi.it/

Versionamento

http://git-scm.com/

http://git.phc.unipi.it/

EOF

Grazie dell'attenzione.