La Matematica nella Società e nella Cultura

Rivista dell'Unione Matematica Italiana

Marco Formentin

Due problemi di meccanica statistica: ricostruzione per catene di Markov su alberi di Galton-Watson e propagazione uniforme del caos e teorema delle fluttuazioni in alcuni modelli di spin

La Matematica nella Società e nella Cultura. Rivista dell'Unione Matematica Italiana, Serie 1, Vol. **3** (2010), n.1 (Fascicolo Tesi di Dottorato), p. 31–34.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=RIUMI_2010_1_3_1_31_0>

La Matematica nella Società e nella Cultura. Rivista dell'Unione Matematica Italiana, Unione Matematica Italiana, 2010.

La Matematica nella Società e nella Cultura
Rivista dell'Unione Matematica Italiana
Serie I, Vol. III, Aprile 2010, 31-34

Due problemi di meccanica statistica: ricostruzione per catene di Markov su alberi di Galton-Watson e propagazione uniforme del caos e teorema delle fluttuazioni in alcuni modelli di spin

Marco Formentin

Un approccio rigoroso a questioni di Fisica Statistica spesso produce interessanti problemi matematici. Questa tesi di dottorato è composta da due parti che non si intersecano, ma entrambe stanno sul confine tra la Teoria della Probabilità e la Meccanica Statistica.

La prima parte tratta il problema della ricostruzione per catene di Markov sugli alberi. Esso ha applicazioni in Informatica e in Biologia. In teoria dell'informazione è equivalente a determinare la capacità critica di un canale rumoroso ramificato, ovvero la quantità di dati che può essere trasmessa in modo "affidabile" attraverso di esso; mentre viene usato in biologia per la ricostruzione filogenetica. In Meccanica Statistica, la non risolubilità del problema della ricostruzione equivale all'estremalità della misura libera di Gibbs sull'albero, caratteristica che dà luogo ad una "nuova" transizione di fase, non osservabile sui reticoli, dove il passaggio estremalità/non-estremalità coincide con la transizione di fase ferromagnetica. Ricordiamo che le misure di Gibbs estremali corrispondono agli stati macroscopici osservabili.
Il problema della ricostruzione può essere descritto come segue. Consideriamo un albero k-regolare (con k figli per ogni vertice) a N livelli. Mandiamo un segnale, scelto casualmente da un alfabeto a q simboli, dall'origine 0 alla frontiera (N-esimo livello) dell'albero. La trasmissione è "rumorosa", nel senso che si ha una certa probabilità di commettere un errore e il simbolo partito dal vertice-padre può essere diverso da quello che giunge al vertice-figlio. Ci si chiede quali informazioni possiamo ricavare sul segnale partito dall'origine conoscendo la configurazione dei simboli alla frontiera dell'albero, quando N tende all'infinito.
Più formalmente, prendiamo un albero T, infinito e senza foglie, e definiamo T^N il sottoalbero formato dall'insieme dei vertici V_N di T con distanza $\leq N$ dall'origine. Per $v, w \in T$, scriveremo $v \to w$ se w è figlio di v.

Ad ogni vertice v associamo una variabile aleatoria $\sigma(v)$ - detta spin - a valori nell'insieme $E = \{1, 2, \ldots, q\}$ e consideriamo la matrice di transizione

$$M = M(v, w)_{1 \leq v, w \leq q} > 0,$$

con distribuzione invariante $a(i)$, per $i = 1, \ldots, q$.

Assegnamo ad ogni configurazione $\sigma_{T^N} \in E^{|V_N|}$ la probabilità

$$\mathbb{P}_N(\sigma_{T^N}) = a(\sigma(0)) \times \prod_{v,w: v \to w} M(\sigma(v), \sigma(w));$$

cioè, il simbolo iniziale viene scelto in accordo con la distribuzione invariante e poi, in ogni ramo, si trasmette da padre a figlio secondo l'evoluzione di una catena di Markov con matrice di transizione M (che in generale dipenderà da alcuni parametri). Definita ξ la configurazione dei simboli al livello N $(=: \partial T^N)$, la grandezza di interesse per leggere la possibilità o meno di ricostruire il segnale originale è il limite della distanza in variazione totale

$$\lim_{N\to\infty} \frac{1}{2} \sum_{\xi \in \partial T^N} |\mathbb{P}_N(\xi|\sigma(0) = i) - \mathbb{P}_N(\xi|\sigma(0) = j)|. \tag{1}$$

Il problema della ricostruzione si dice risolubile se (1) > 0 per almeno una coppia di i, j in E. Altrimenti, quando la distanza in variazione totale è sempre nulla al limite, si ha non risolubilità, equivalente al fatto che non c'è trasporto di informazione dall'origine alla frontiera dell'albero per un segnale tipico. La possibilità di ricostruire il segnale dipende dall'errore che si commette ad ogni passo, ovvero dalla matrice M, da q e da k.

Una condizione sufficiente per la ricostruzione è dovuta a Kesten e Stigum. Essi hanno dimostrato che dato un albero k-regolare e detto λ il secondo autovalore più grande in modulo della matrice M, c'è ricostruzione se $k\lambda^2 > 1$. Quest'ultima condizione (nel seguito KS), in generale non è esatta, cioè non è anche necessaria. Negli ultimi anni la ricerca di un bound esatto per la ricostruzione ha impegnato molti ricercatori. A tal proposito, l'esempio più studiato in letteratura è il modello di Potts, dove la matrice di transizione M è simmetrica. In questo caso la condizione KS si sa essere necessaria e sufficiente solo nel caso $q = 2$ (modello di Ising). Per $q \geq 3$ la determinazione della soglia esatta della transizione ricostruzione/non-ricostruzione è ancora un problema aperto. Esiste una congettura. Nel 2006, sulla base di simulazioni numeriche, Montanari e Mézard hanno ipotizzato che per il modello di Potts il bound KS sia esatto per $q \leq 4$, mentre non lo sia mai quando $q \geq 5$. Questa congettura è stata parzialmente provata nel 2008 quando Sly dimostra che KS è esatto per $q = 3$ e k abbastanza grande e non lo è mai se $q \geq 5$. L'ipotesi su k è necessaria per poter utilizzare un'approssimazione di tipo Teorema del Limite Centrale. Il risultato non si applica ad alberi regolari con k piccolo e non dà un'espressione per la soglia di non-ricostruzione quando $q \geq 5$. Manca il caso $q = 4$. Nella tesi dimostriamo una condizione sufficiente per la non-ricostruzione che, per k piccolo e $q = 3$ o 4, discorda da KS dalla terza cifra decimale in poi e che migliora le precedenti stime sufficienti presenti in letteratura. Quando $q > 4$ l'espressione differisce di poco dai valori calcolati numericamente da Montanari e Mézard.

Per enunciare il teorema, detto p un vettore di probabilità in E e a la distribuzione invariante per M, chiamiamo entropia simmetrica la funzione

$$L(p) = S(a|p) + S(p|a) = (p-a)\log\left(\frac{dp}{da}\right) = \sum_{i=1}^{q} (p(i) - a(i)) \log \frac{p(i)}{a(i)}.$$

La nostra condizione [3], [4] dice che su un albero k-regolare non c'è ricostruzione se

$kc(M) < 1$, con

$$c(M) = \sup_p \frac{L(pM^{\text{rev}})}{L(p)}$$

costante che dipende dalla matrice di transizione, poiché $M^{\text{rev}}(i,j) = \dfrac{M(j,i)a(j)}{a(i)}$. Il metodo usato nella dimostrazione è alquanto generale e permette di trattare un'ampia classe di matrici di transizione. Per catene simmetriche 2-dimensionali la condizione data è necessaria e sufficiente per la ricostruzione. In generale il bound non è esatto, ma migliora le stime di Peres e Mossel del 2003, sia per il modello di Potts, che per le catene "fortemente" asimmetriche. Tutti i risultati qui citati si generalizzano senza difficoltà ad alberi di Galton-Watson.

Nella seconda parte della tesi studiamo la propagazione uniforme del caos. La propagazione del caos è un fenomeno noto per processi stocastici con interazione a campo medio [1]. Con processo stocastico a campo medio intendiamo una famiglia $x^{(N)} = (x^{(N)}(t))_{t\geq 0}$, dove $x^{(N)}(t) = (x_1^{(N)}(t), x_2^{(N)}(t), \ldots, x_N^{(N)}(t))$ è un processo di Markov con N componenti reali. Inoltre, se definiamo la misura empirica

$$\rho_N(t) := \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k^{(N)}(t)},$$

allora $(\rho_N(t))_{t\geq 0}$ è anch'essa un processo di Markov.

Supponiamo che $X^{(N)}(0)$ abbia legge prodotto con marginali λ, diciamo che vale la propagazione del caos se

1. per ogni $t > 0$, $\rho_N(t)$ converge in probabilità ad una misura deterministica $\rho(t)$, quando $N \uparrow \infty$;
2. siano $i_1, i_2, \ldots, i_m$ degli indici fissati e $T > 0$, allora il processo stocastico

$$\big(x_{i_1}^{(N)}(t), x_{i_2}^{(N)}(t), \ldots, x_{i_m}^{(N)}(t)\big)_{t\in[0,T]} \tag{2}$$

converge in legge, quando $N \uparrow \infty$, al processo $(\xi_1(t), \ldots, \xi_m(t))_{t\in[0,T]}$ a componenti i.i.d. con legge $\rho(t)$.

Inoltre, diciamo che vale un "teorema delle fluttuazioni" se per ogni $h : \mathbb{R} \to \mathbb{R}$, funzione continua e limitata, la fluttuazione

$$\left(\sqrt{N}\left[\int h d\rho_N(t) - \int h d\rho(t)\right]\right)_{t\in[0,T]} \tag{3}$$

converge debolmente ad un processo Gaussiano.

Spesso i modelli a campo medio dipendono da alcuni parametri e possono avere delle transizioni di fase. Noi consideriamo transizioni di fase descritte in termini del comportamento limite di $\rho(t)$ che diremo subcritico se $\rho(t)$ converge a $\rho(\infty)$ indipendentemente dalle condizioni iniziali. Supercritico altrimenti.

Lo scopo è quello di indagare l'uniformità nel tempo della propagazione del caos e del teorema delle fluttuazioni per alcuni di questi modelli. Più precisamente:

- La convergenza $(x_{i_1}^{(N)}(t), \ldots, x_{i_m}^{(N)}(t))$ a $(\xi_1(t), \ldots, \xi_m(t))$ è uniforme in t?
- La convergenza delle fluttuazioni $\left(\sqrt{N}\left[\int h d\rho_N(t) - \int h d\rho(t)\right]\right)$ verso il limite Gaussiano è uniforme in t?

In questa tesi riusciamo a provare l'uniformità nel tempo dei limiti di (2) e (3) per due sistemi di spin a campo medio nel caso sottocritico. Nel regime supercritico l'uniformità è impossibile nel senso della definizione che abbiamo dato. Il primo è il modello di Curie-Weiss, che può essere considerato come il più semplice modello a campo medio. Il secondo modello è stato proposto nel contesto del rischio di credito [2] per un network di aziende legate da relazioni d'affari. Dal punto di vista matematico il modello non è reversibile e la misura invariante non è conosciuta. Le seguenti considerazioni giustificano lo studio dell'uniformità dei teoremi limite. Molte quantità di interesse finanziario sono basate sull'approssimazione normale di medie empiriche. Nelle applicazioni finanziarie N è dato ed è dell'ordine di qualche centinaio. L'uniformità nel tempo esclude che l'approssimazione si deteriori col tempo e diventi irrealistica per tempi lunghi. Inoltre, la propagazione del caos e il teorema delle fluttuazioni uniformi danno un CLT e una LLN per la media empirica rispetto alla misura invariante del sistema che può non essere nota. Perciò risultati sulla statica possono essere derivati dallo studio della dinamica.

Anche se abbiamo scelto di lavorare con modelli specifici, crediamo che i metodi impiegati si possano usare anche per altre classi di modelli.

BIBLIOGRAFIA

[1] CATTIAUX P., GUILLIN A. e MALRIEU F., *Probabilistic approach for granular media equations in the non-uniformly convex case*, Probability Theory and Related Fields, **140** (2008), 19-40.

[2] DAI PRA P., RUNGGALDIER W. J., SARTORI E., e TOLOTTI M., *Large Portfolio Losses: A dynamic contagion model*, The Annals of Applied Probability, **19** (2009), 347-394.

[3] FORMENTIN M. e KÜLSKE C., *On the Purity of the free boundary condition Potts measure on random trees*, Stochastic Processes and their Applications, **119** (2009), 2992-3005

[4] FORMENTIN M. e KÜLSKE C., *A symmetric entropy bound on the non-reconstruction regime of Markov chains on Galton-Watson trees*, Electronic Communications in Probability, **14** (2009), 587-596.

Fakultät für Mathematik, Ruhr-Universität Bochum
e-mail: marco.formentin@rub.de
Dottorato in Matematica con sede presso l'Università di Padova – Ciclo XXI
Direttori di ricerca: Prof. Dr. Christof Külske, Ruhr-Universität Bochum
e Prof. Paolo Dai Pra, Università di Padova