Julián Martínez



instituto de cálculo

UBA - CONICET



Conicet - IC / Departamento de Matemática - FIUBA

Projectos presentes

Mecánica Estadística - Procesos Estocásticos
Límite Hidrodinámico - KPP





Matthieu Jonckheere - IC
Pablo Groisman - DM

Variational autoencoders - ML
Procesos gaussianos





Pavlos Protopapas - Harvard
Javier Kreiner

Embbedings - ML
Word2vec dinámico







Ignacio Hamelin-FIUBA
Mariano Beiro-FIUBA
Carlos Selmo-ITBA

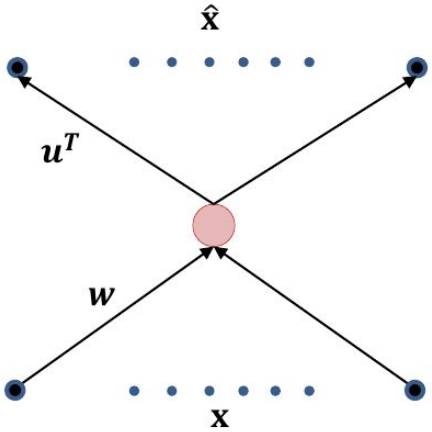
Procesos estocásticos - economía
Equilibrio general

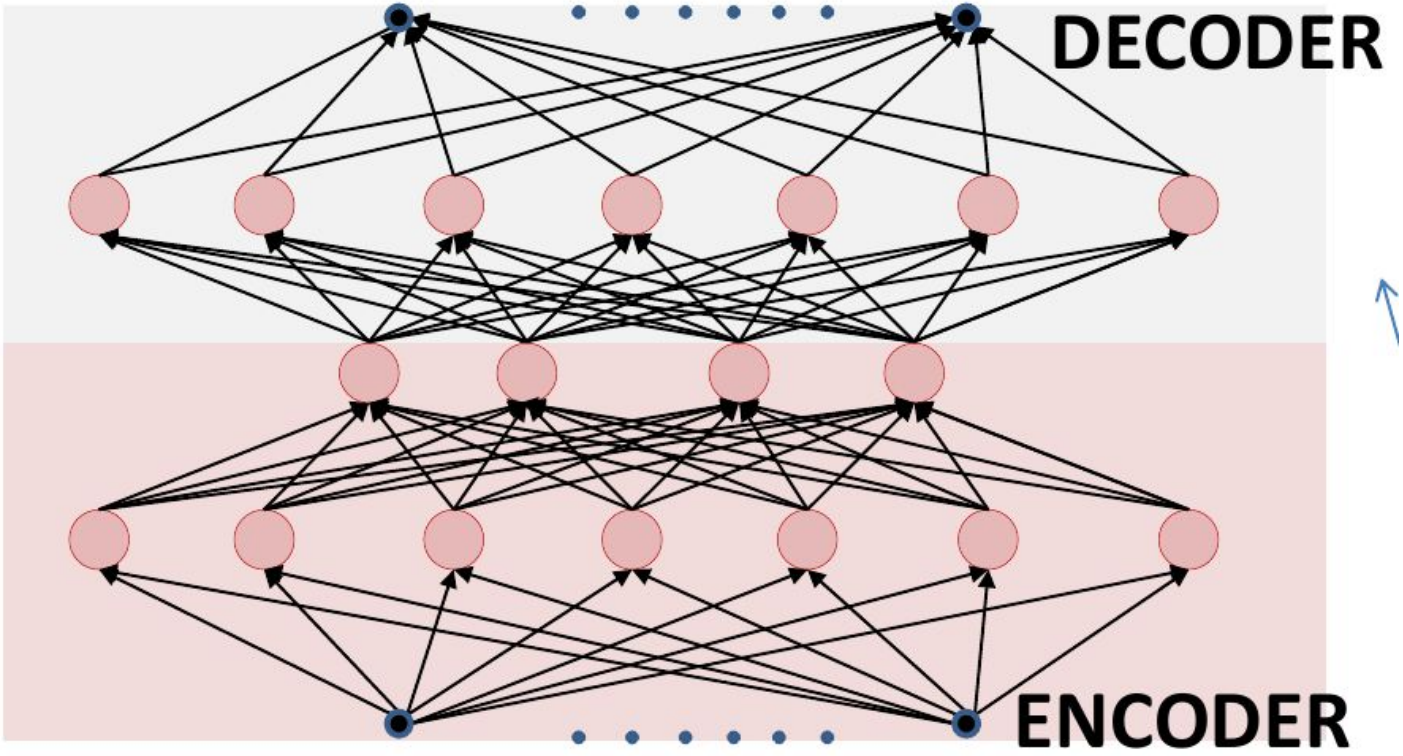




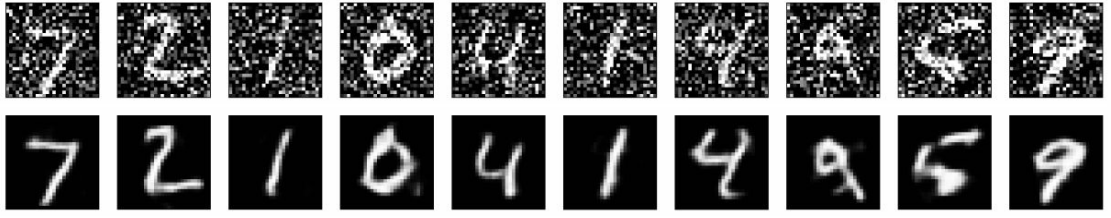
Damian Pierri - IEP
Nicolas Frevenza - UDELAR

Autoencoders



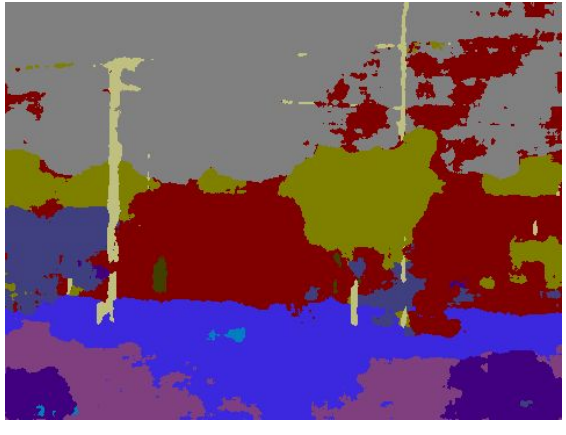


Autoencoders-usos

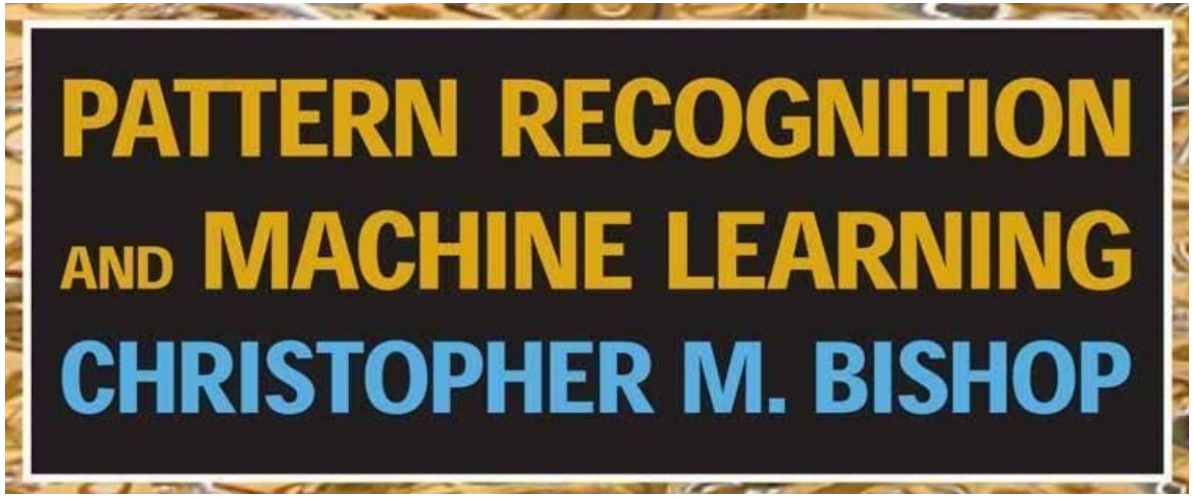


- Quitar ruido de datos
- Encontrar features relevantes
- Clustering
- segmentación





Sobre la notación



???

$$\mathrm{H}[x] = -\sum_{x} p(x) \log_2 p(x). \tag{1.93}$$

Variational Autoencoders - (Versión Probabilística)

$$Z \sim N(0, I_2) \quad ; \quad X \mid Z = z \sim p(\cdot \mid z)$$

Quiero hacer inferencia

$$p(z \mid x) = \frac{p(x \mid z)\, p_Z(z)}{p_X(x)}$$

$$p_X(x) = \int p(x \mid z)\, p_Z(z)\, dz$$

evidencia

Inferencia Variacional

$$Z \mid X = x \sim q_\lambda(\cdot \mid x)$$

aproximo por una familia más simple

$$J(\lambda) = KL\left(q_\lambda(\cdot \mid x) \,\|\, p(\cdot \mid x)\right)$$

$$KL(q \,\|\, p) = E_q[\log q(Z)] - E_q[\log p(Z)]$$

Otro enfoque



$$X = f(Z) + \varepsilon \qquad Z \sim \mathcal{N}(0, I_2)$$

$$\varepsilon \sim \mathcal{N}(0, I_N)$$

$$p_\theta^d(x|z) \sim \mathcal{N}\left(f_\theta^d(z), I_N\right)$$

Variational Autoencoders - Esquema

ENCODER

$$q_\lambda^e(z|x) \sim \mathcal{N}\left(\mu_\lambda^e(x) ; \sigma_\lambda^e(x)^2 I_2\right)$$

DECODER

$$p_\theta^d(x|z) \sim \mathcal{N}\left(f_\theta^d(z) ; I_N\right)$$

Variational Autoencoders - ELBO

$$KL\left(q_\lambda(\cdot|x) \,\|\, P(\cdot|x)\right) = E_{q_\lambda}\left[\ln q_\lambda(z|x)\right] - E_{q_\lambda}\left[\ln p(z|x)\right]$$

$$= \underbrace{KL\left(q_\lambda(\cdot|x) \,\|\, P_z\right) - E_{q_\lambda}\left[\ln p(x|z)\right]}_{\text{ELBO}} + \ln P_x(x)$$

maximizando obtengo una cota inferior para $\ln P_x(x)$

Variational Autoencoders - funcion de costo

$$\Psi(\lambda,\theta) = E_{q^e_\lambda}\left[\log p^d_\theta(x|z)\right] - KL\left(q^e_\lambda(\cdot|x)\,\|\,P_z\right)$$

Reconstruir

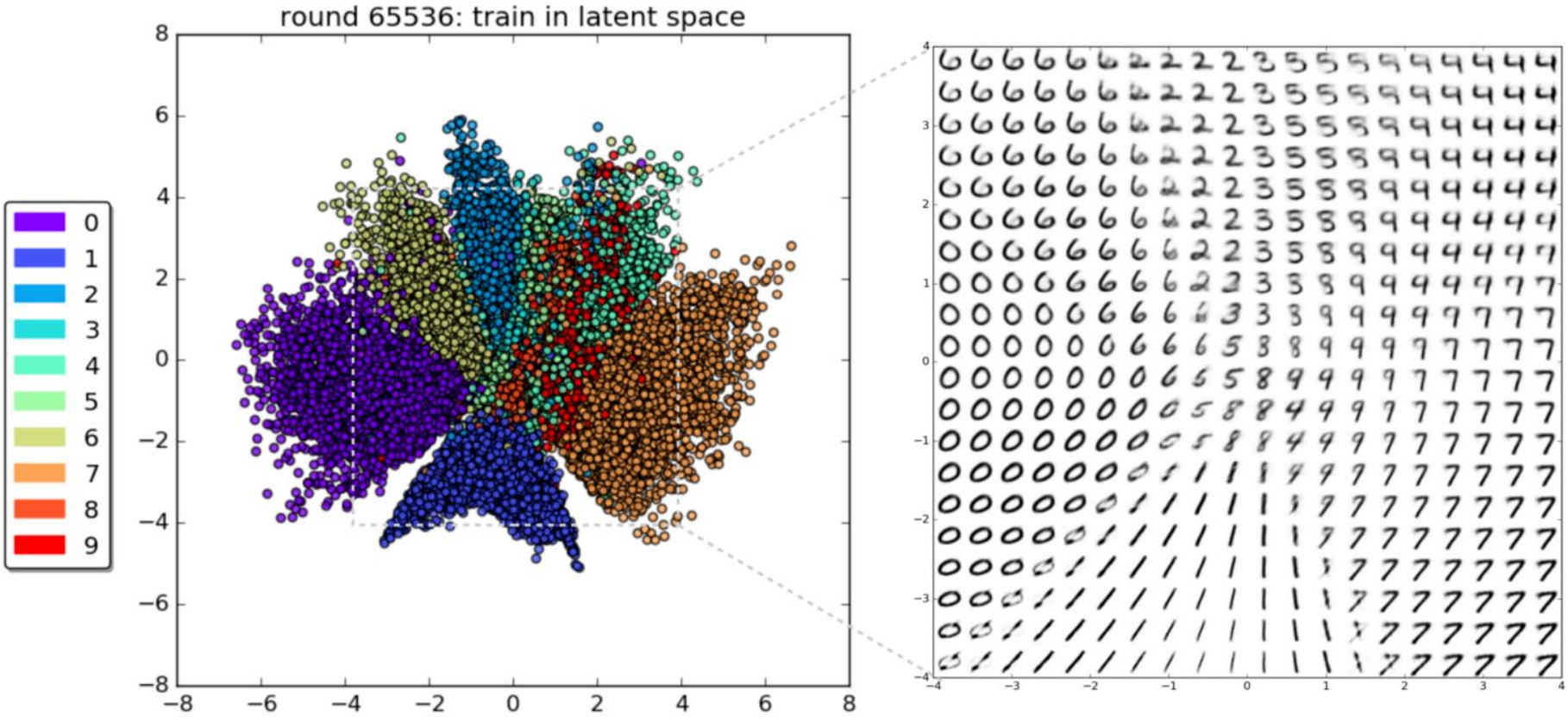
regularización

reparametrization trick

$$z = \mu^e_\lambda(x) + W\sigma^e_\lambda(x)$$

$$W \sim \mathcal{N}(0, I_z)$$

$$x \longrightarrow \left(\mu_\lambda^e(x), \sigma_\lambda^e(x)\right) \longrightarrow f_\theta^d\left(\mu_\lambda^e(x) + W\sigma_\lambda^e(x)\right) \rightarrow \|x - f_\theta^d(\dots)\|^2$$



La idea

Auto-encoding **variational** bayes

DP Kingma, M Welling - arXiv preprint arXiv:1312.6114, 2013 - arxiv.org

… Regular- ization techniques have been proposed to make **autoencoders** learn useful representations … objective contains a regularization term dictated by the **variational** bound (eg … recently introduced Generative Stochastic Networks [BTL13] where noisy **auto-encoders** learn …

☆ ❞ Citado por 4211 Artículos relacionados Las 23 versiones »



Variational Autoencoders - claves

$$\mathbb{E}\left[\log p_\theta^d\left(x \mid \mu_\lambda^e(x) + W\sigma_\lambda^e(x)\right)\right] - KL\left(q_\lambda^e(\cdot|x) \| p_z\right)$$

Reparametrization Trick

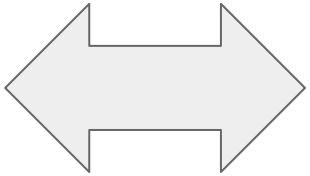
Explícita

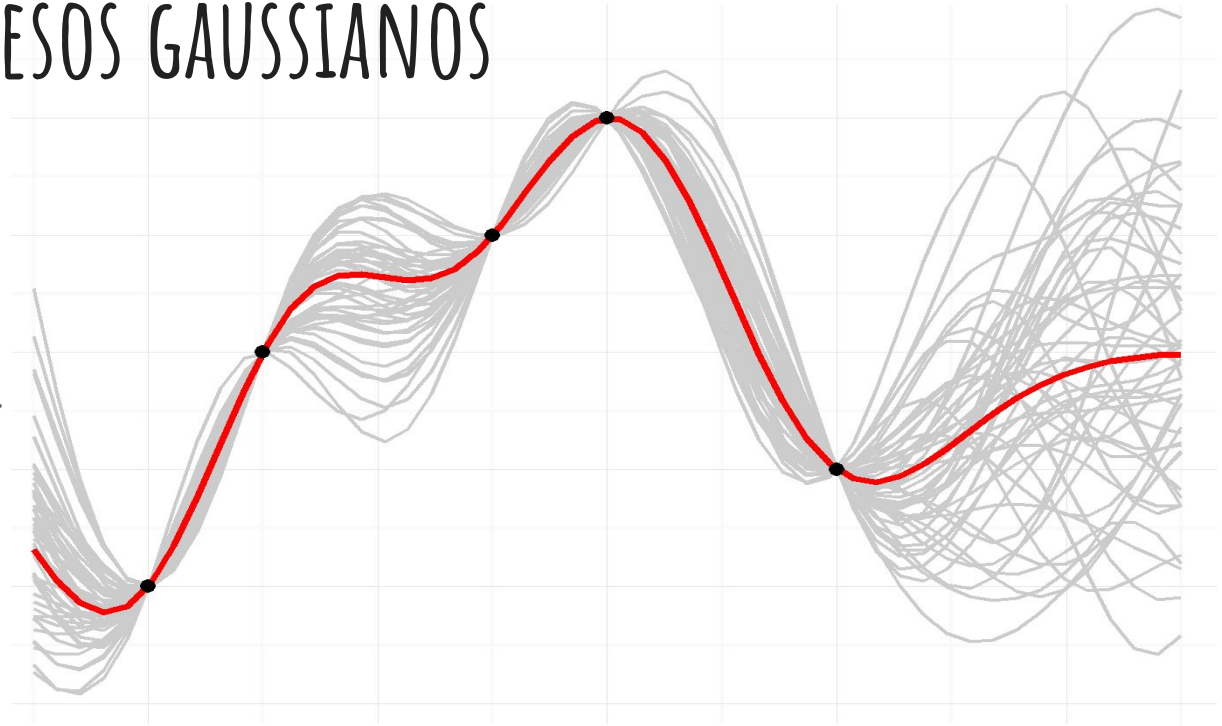
¿Con qué otros modelos puedo hacer lo mismo?

¿Otros espacios latentes pueden que funcionen mejor?

La propuesta - procesos gaussianos







$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(\mathbf{x}_p - \mathbf{x}_q)^2\right) + \sigma_n^2 \delta_{pq}$$



Docencia / Cursos

- Probabilidad y Estadística - FIUBA
- Aprendizaje Estadístico - Optativa de Doctorado - FIUBA (junto con Jemina García)
- Variational Inference - Minicurso en la Universidad Federal de Bahía
- Aprendizaje Reforzado - Maestría en Exploración de datos y Descubrimiento del conocimiento (junto con Javier Kreiner)

Posibles ideas

Variational Inference: A Review for Statisticians

David M. Blei
Department of Computer Science and Statistics
Columbia University

5.4 Open problems

There are many open avenues for statistical research in variational inference.

Posibles ideas

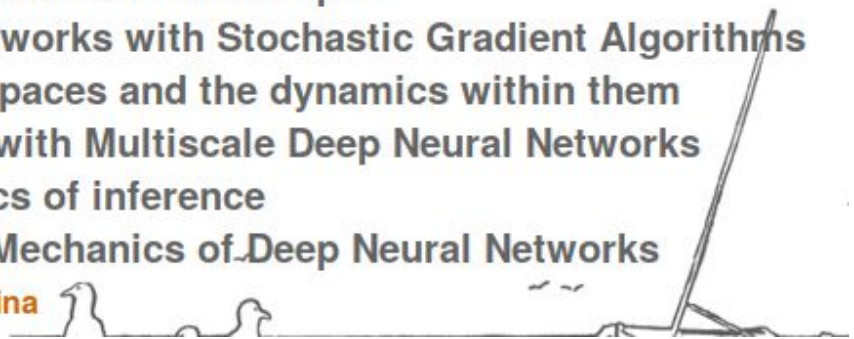
The Beg Rohu Summer School

Home Aims Information Sailing History Apply

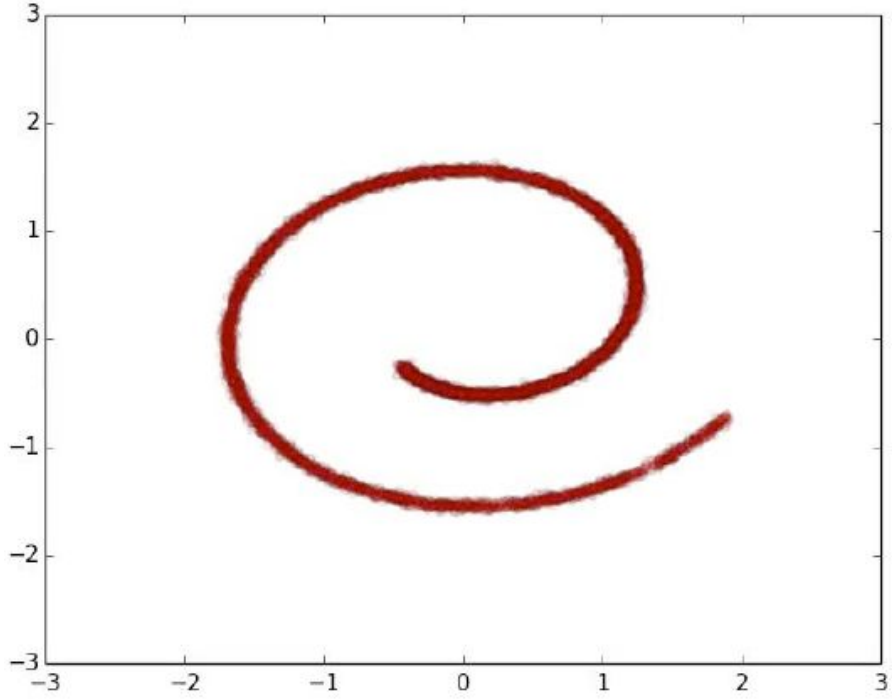
25 June - 7 July 2018

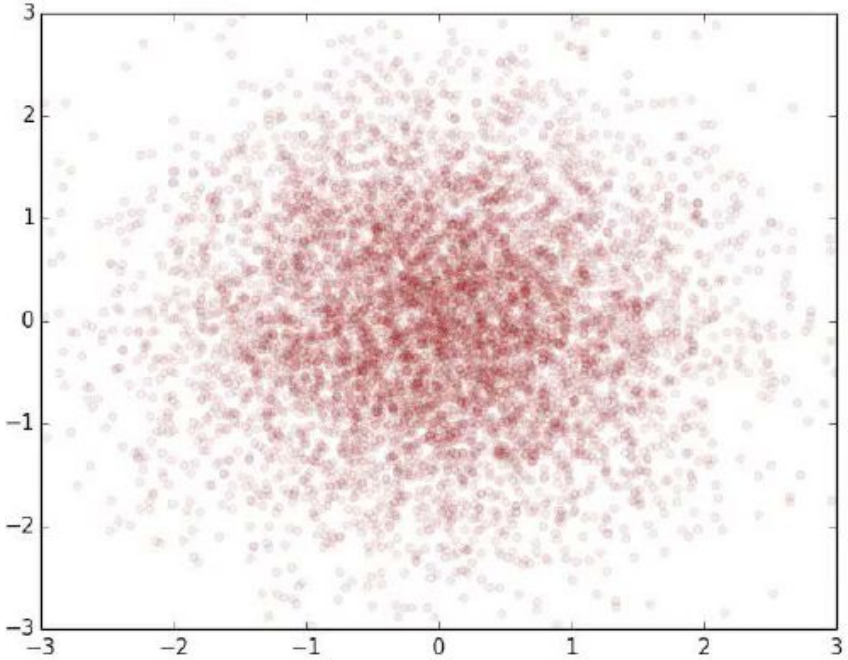
Deep Learning and Statistical Physics

G. Ben Arous: Complexity of random landscapes
L. Bottou: Training Deep Networks with Stochastic Gradient Algorithms
J. Kurchan: Rugged phase-spaces and the dynamics within them
S. Mallat: Learning Physics with Multiscale Deep Neural Networks
M. Mézard: Statistical physics of inference
H. Sompolinsky: Statistical Mechanics of Deep Neural Networks
Seminars: **F. Krzakala and R. Zecchina**

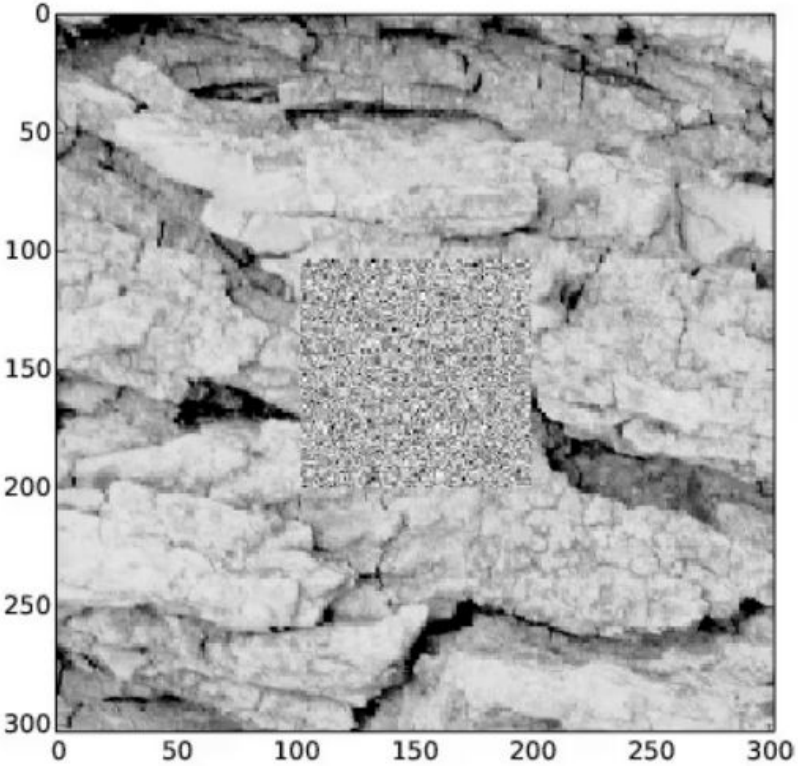


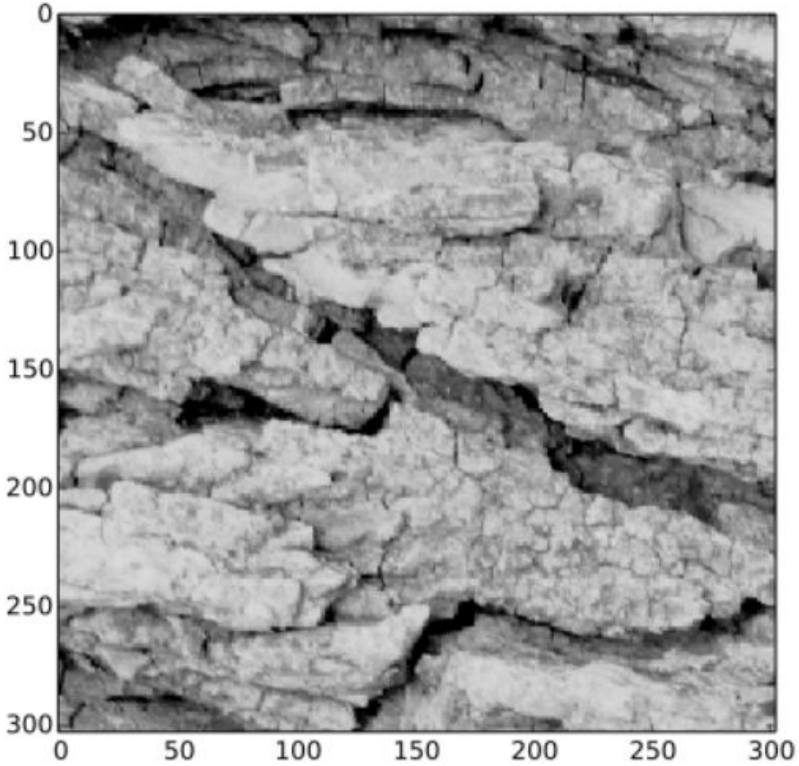
Desordenando (Fuera de equilibrio)





Modelling arbitrary probability distributions using non-equilibrium thermodynamics,
J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, ICML 2015.





Datos de color

CS 174
Combinatorics and Discrete Probability

Prof. Michael Jordan
Monday, Wednesday and Friday, 1:00-2:00, 60 Evans Hall
Spring 2015

Announcements

- Announcements will be put on the piazza website: piazza.com

---

Topics

- Random variables, mean, variance
- Tail bounds: Markov, Chebyshev, Chernoff/Hoeffding
- Balls, bins and permutations
- Random graphs
- The probabilistic method
- The Lovasz local lemma
- Markov chains
- Entropy, information and coding
- Markov chain Monte Carlo
- Coupling
- Martingales
