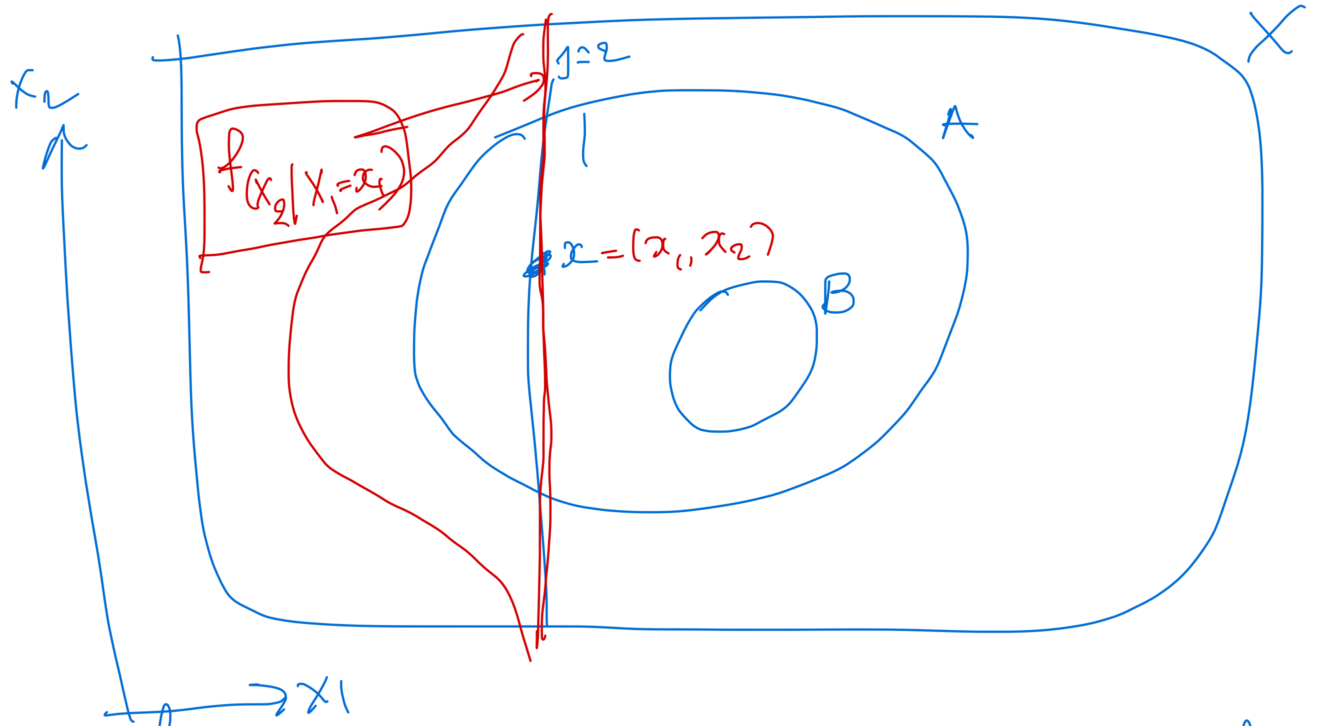
17-5-2023

Εφαρμογή Gibbs Sampler

$X \in \mathbb{R}^n$ $\quad f(\underline{x})$, υπάρχουν γεννήτριες για τις $f_{X_j|X_{-j}}(x_j|x_{-j})$

$$\theta = P(X \in B \mid X \in A), \quad B \subseteq A$$



$n=2$

$f_{X|A}$ → Generate N obs from $f_{X|A}$

$X_1, \dots, X_N$

$$\hat{\theta} = \sum_{i=1}^{N} 1(X_i \in B) \Big/ N$$

① Αλγόριθμος Gibbs για παρατηρήσεων από $f_{X|A}(x|X \in A)$

$$f_{X|A}(x) = \begin{cases} \dfrac{f(x)}{P(X\in A)}, & x \in A \\ 0, & \text{αλλού} \end{cases} = C\, f(x), \quad x\in A$$

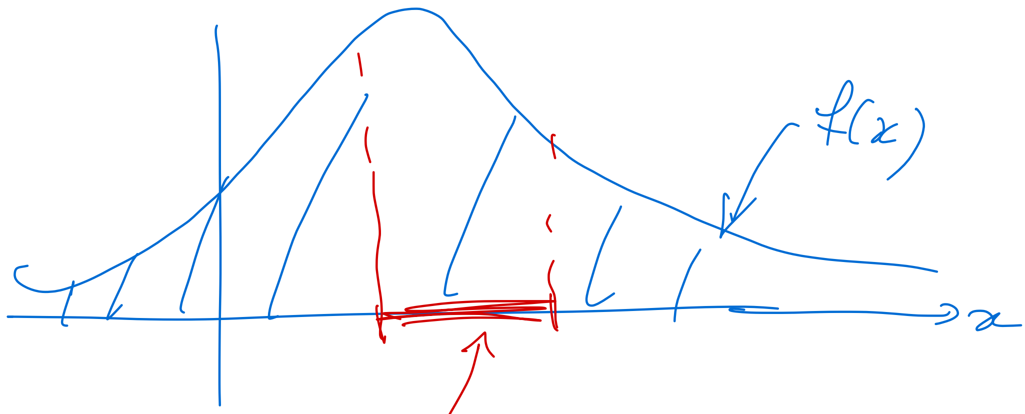
$$= C \underbrace{1(x\in A)\, f(x)}_{b(x)}$$

Gibbs sampler

$x_0 \in A \quad (x_0 \in \mathbb{R}^n)$

<u>Υποπρόβλημα</u> Έστω $X \in \mathbb{R}$ $f_X(x)$: ∃ γεννήτρια

θέλουμε γεννήτρια από $f_{X|X\in M}$ $(M \subseteq \mathbb{R})$



Απλή

<u>Accept/Reject</u>

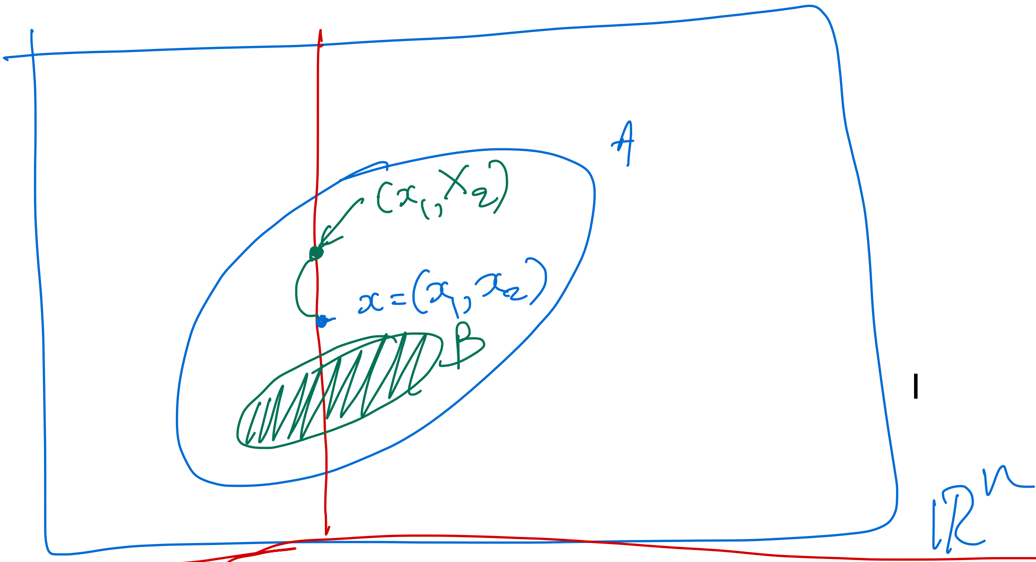
① Δημ $Y \sim f$

Αν $Y \in M \Rightarrow X = Y$

διαφ. απορρ. επιστρ. ①

Πίσω στο αρχικό πρόβλημα

γεννήτρια Gibbs από $f_{X|X\in A}$, $X\in\mathbb{R}^n$



αρχ. κατάσταση

Έστω $x\in A$, $x=(x_1,\dots,x_n)$

Επιλ. τυχαία διεύθυνση $j\in\{1,2,\dots,n\}$

Έστω $j=2$

Θέλουμε παρατήρηση από $f_{\cdot}(x_2 \mid X_1=x_1, X_2\in A)$

① Δημ. $Y\sim f_{X_2|X_1=x_1}$

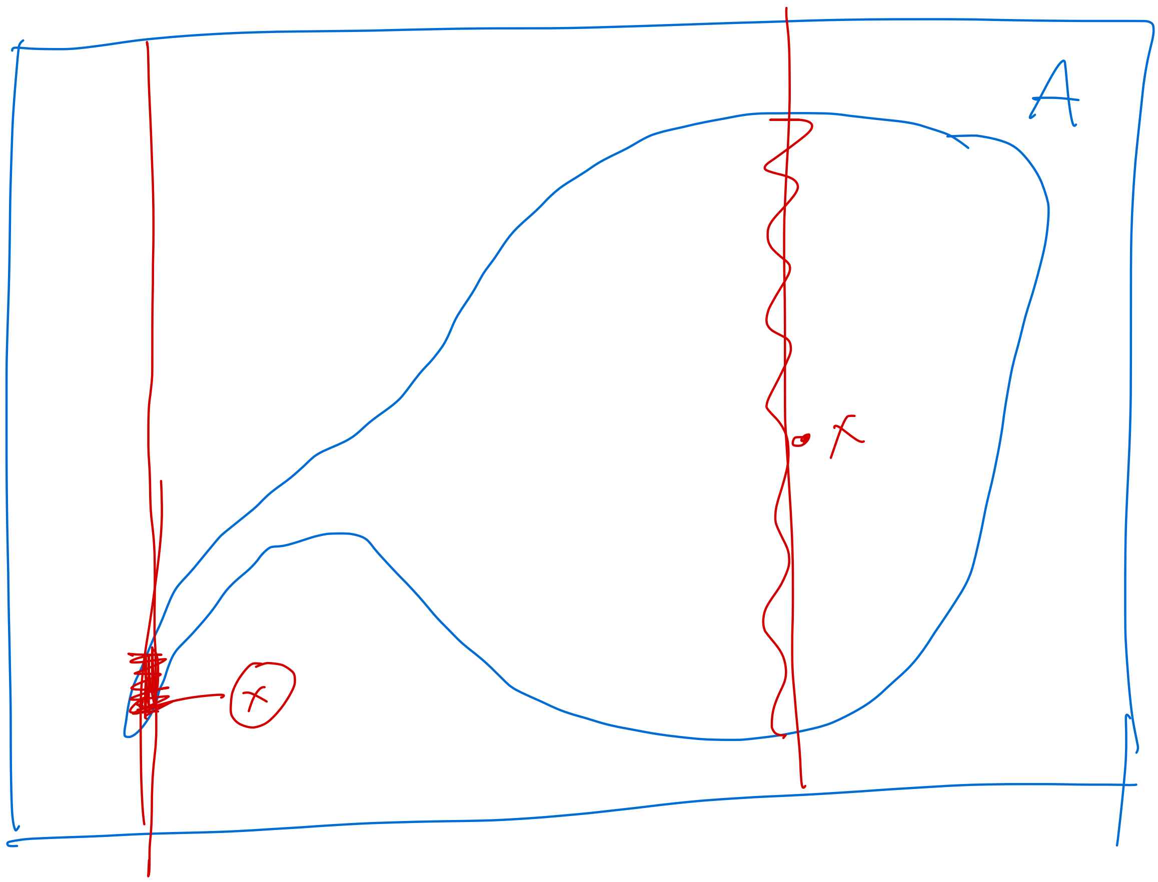
② Αν $Y\in A \Rightarrow X_2=Y$

Αν $Y\notin A \Rightarrow$ απορ. επιστρ. ①

$\Rightarrow X_2$

Νέα κατάσταση: (x_1, X_2) [Gibbs δεκτή]

Gibbs sampler για $f_{X|X\in A}$



A
x
x

Τώρα έχουμε γεννήτρια Gibbs για $f_{X|X\in A}(x|x\in A)$

Δηλ. $\{X_1, X_2, \dots\}$ Markov Chain

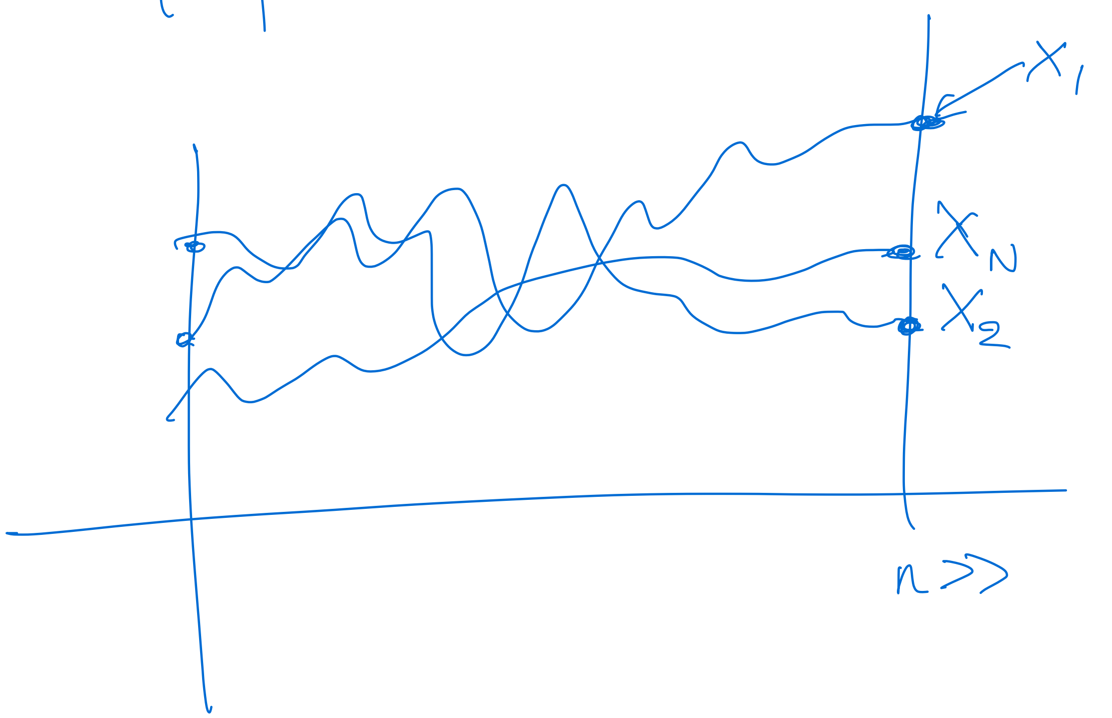
με στάσιμη κατανομή $\pi(x) = f_{X|X\in A}(x)$

Θέλουμε $\theta = P(X\in B \mid X\in A)$

<u>1η προσέγγιση</u> Δημιουργούμε N παρατ. από $\pi(x)$

$$\hat{\theta} = \frac{1}{N}\sum_{j=1}^{N} 1(X_j \in B)$$

<u>"Πρόβλημα"</u> Κάθε παρατήρηση $X_j \sim \pi$ απαιτεί προσομοίωση της $\{X_n\}$ MC για $n \gg$

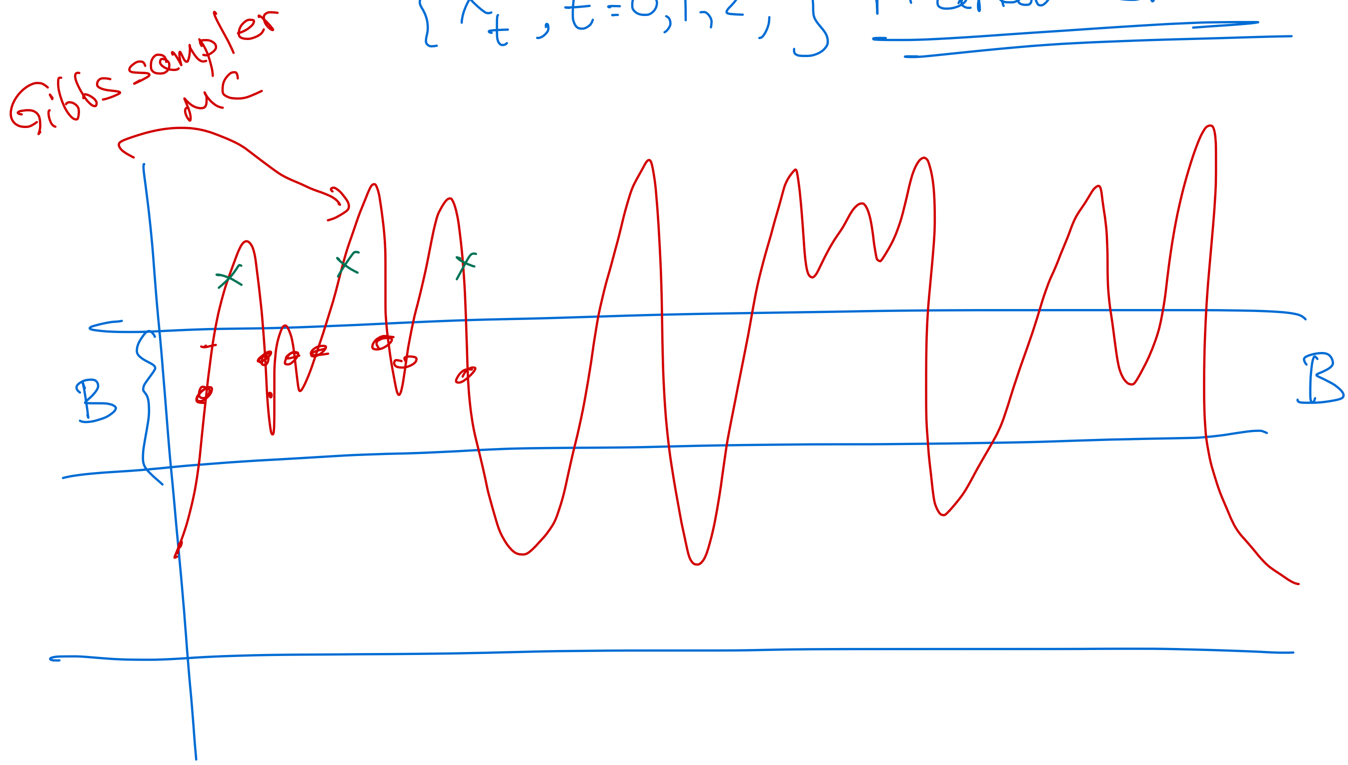


2η προσέγγιση

$$\pi_j = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} 1(X_t = j)$$

$$\boxed{P_\pi(X\in B) = E_\pi(\underbrace{1(X\in B)}_{c(X)}) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} 1(X_t\in B)}$$

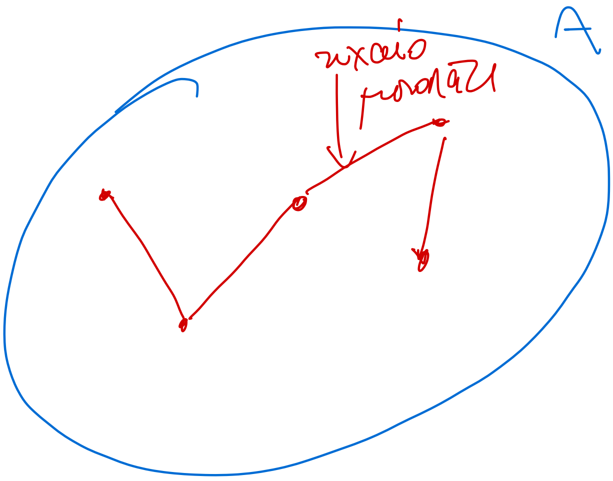
$\{X_t, t=0,1,2,\}$ Markov Chain



Simulated Annealing (Stochastic Optimization)

$$z = \max_{x \in A} f(x) \quad \Big\} \to \text{αλγόριθμοι τυχαιοποίησης}$$

$$M = \{x \in A : f(x) = z\}$$

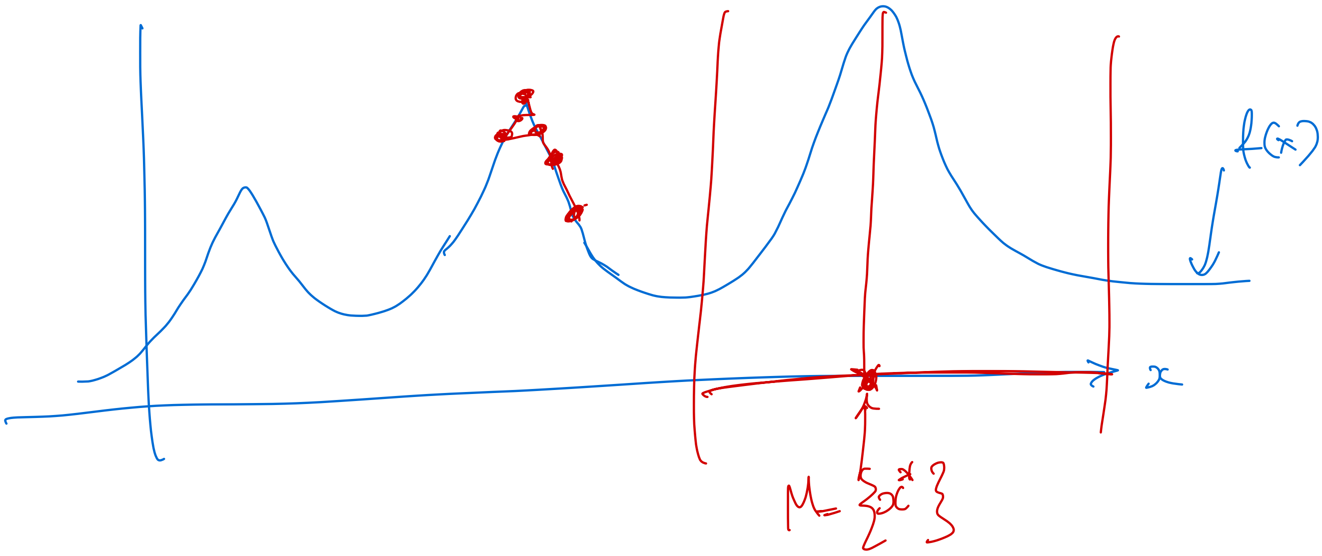


$\{X_1, X_2, \dots\}$ τυχαία ακολουθία $X_j \in A, \forall j$

1) $\lim_{n\to\infty} P(X_n \in M) = 1$ ($X_n \to M$ κατά πιθανότητα)

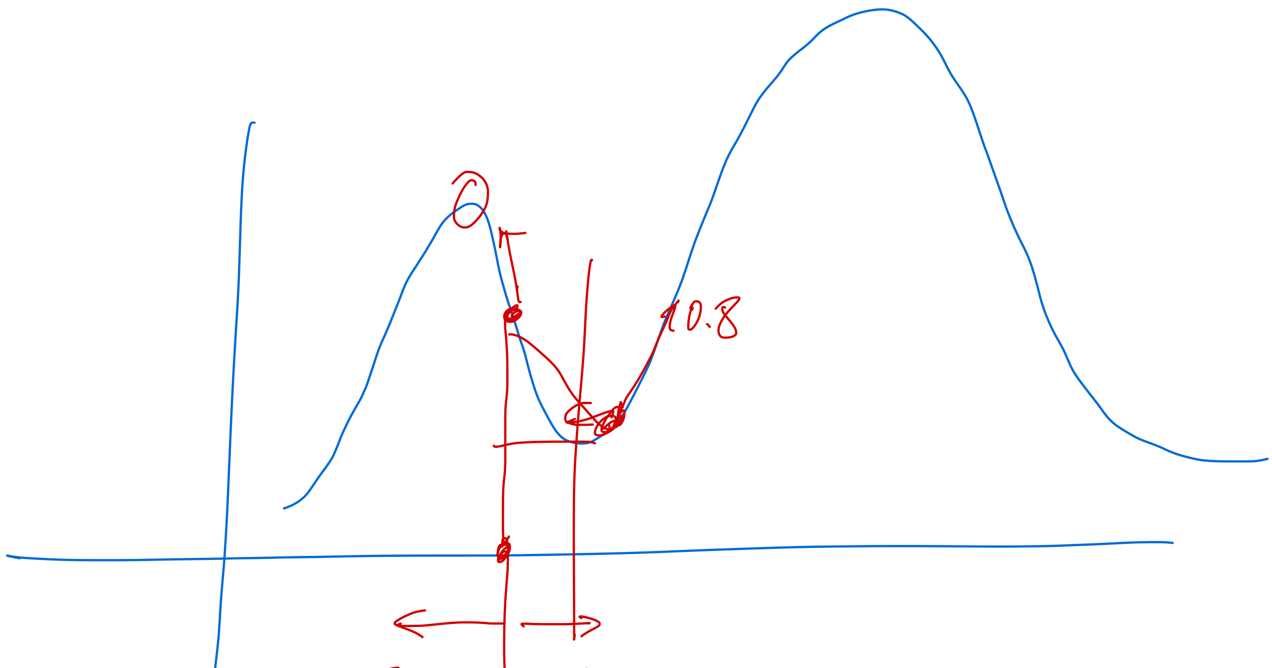
2) $P(\lim_{n\to\infty} X_n \in M) = 1$ ($X_n \to M$ σχεδόν βεβαίως)

1. Simulated Annealing
2. Stochastic Approximation (Stochastic Gradient)



Κινούμαστε πρός κατεύθυνση αύξησης της $f(x)$

Σε κάθε x υπολογίζουμε $f(x)$, $\nabla f(x)$



∇ $-\nabla$

$+$

80%   20%

p_n $1-p_n$

n: βήμα

Markov Chain Monte Carlo Approach.

$A \subseteq \mathbb{R}^n \qquad |A| < \infty \qquad V : A \to \mathbb{R}$

$V(x) \geq 0 \quad \forall x \in A$

$$V^* = \max_{x \in A} \{ V(x) \}$$

$$\mathcal{M} = \{ x \in A : V(x) = V^* \} = \arg\max_{x \in A} \{ V(x) \}$$

$\boxed{V^* = ? \quad , \quad \text{ένα } x \in \mathcal{M}}$ ← Πρόβλημα

Έστω τυχαία μετ. $X \in A$

$$P(X = x) = p_\lambda(x) = \frac{e^{\lambda V(x)}}{\sum_{x \in A} e^{\lambda V(x)}} = C e^{\lambda V(x)}$$

$\lambda > 0$

αν $x_1, x_2 \qquad V(x_1) < V(x_2) \Rightarrow$

$p_\lambda(x_1) < p_\lambda(x_2)$

Όταν $\lambda \to \infty \qquad \lim_{\lambda \to \infty} p_\lambda(x) = ?$

$$p_\lambda(x) = \frac{e^{\lambda V(x)} \, e^{-\lambda V^*}}{e^{-\lambda V^*} \sum_x e^{\lambda V(x)}} = \frac{e^{\lambda (V(x) - V^*)}}{\sum_x e^{\lambda (V(x) - V^*)}} =$$

$$= \frac{e^{\lambda (V(x) - V^*)}}{|M| + \underbrace{\sum_{x \notin M} e^{\overbrace{\lambda (V(x) - V^*)}^{<0}}}} \quad \forall \lambda \qquad \begin{aligned} V(x) - V^* &\le 0 \quad \forall x \\ &= 0, \quad x \in M \end{aligned}$$

Όταν $\lambda \to \infty$

$$\begin{cases} \sum_{x \notin M} e^{\lambda (V(x) - V^*)} \to 0 \\[2ex] \lim_{\lambda \to \infty} e^{\lambda (V(x) - V^*)} = \begin{cases} 1, & x \in M \\ 0, & x \notin M \end{cases} \end{cases}$$

$$\lim_{\lambda \to \infty} P_\lambda(x) = \frac{1(x \in M)}{|M|}$$

διακρ. ομοιόμ. στο M

Ερωτήματα:

① Πώς παίρνουμε παραδειγμ. από $P_\lambda(x)$?

[MCMC]

(Metropolis)

② $\lambda \to \infty$? πώς το ορίζουμε;