Apply Metropolis – Hasting Algo to Optimatization of a function and we will review "simulated annealing".

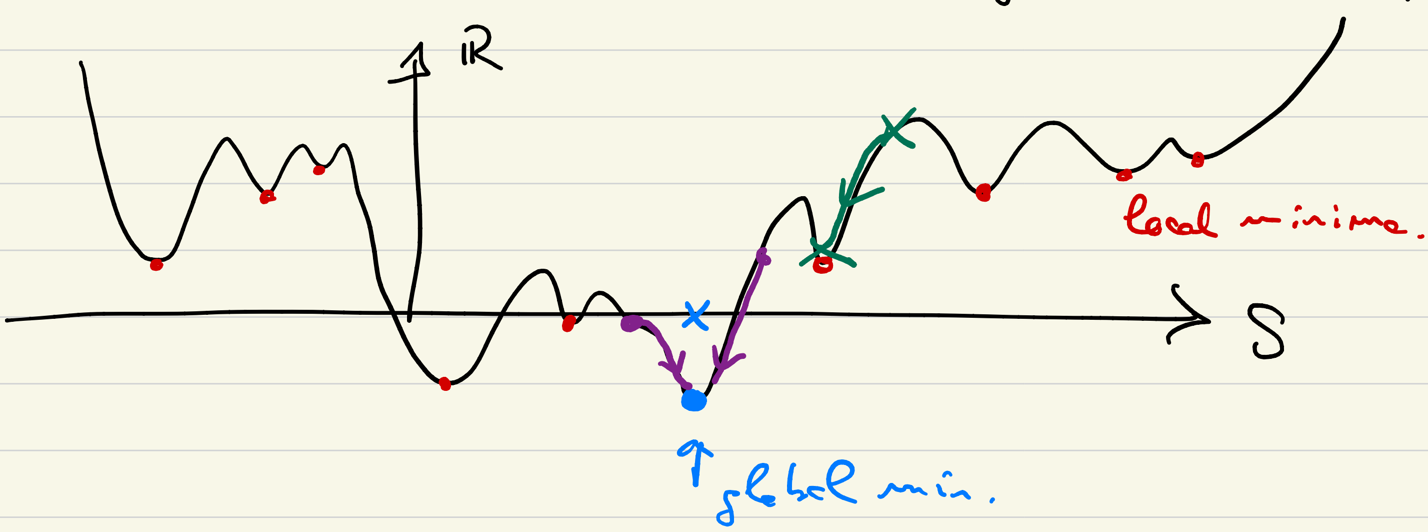
function $f : S \longrightarrow \mathbb{R}$; $i \in S \mapsto f(i)$.

S ↑ discrete space

$\mathbb{Z}$; Colorings of graphs; assignement of binary variables in Ising model

Goal : find a global minimum of f.

(content to find some "good minimum of f")



Examples of functions of interest:

- optimization probl is hard generally for high dimensional fcts.

<u>Coloring</u> $G = (V, E)$. Proper coloring is an assignment of colors $x \in \{1, \dots, q\}$ to (q colors) $k \in V$ s.t. $(k, \ell) \in E$ $x_k \neq x_\ell$.

State space $i \leftrightarrow \underline{x} = (x_1, \dots x_{|V|})$.

Minimize the "<u>cost fct</u>":

$$f \leftrightarrow C(\underline{x}) = \sum_{(k,\ell) \in E} \mathbb{1}(x_k = x_\ell) . \geq 0$$

= Number of edges whose two vertices have the same color.

$$C : S = \{1, \dots, q\}^V \to \mathbb{R},$$

$$\underline{x} \mapsto C(\underline{x}).$$

Ising Hamiltonian or energy fct or cost fct:

$G = (V, E)$. assign $S_k = \pm 1$ to vertices $k \in V$

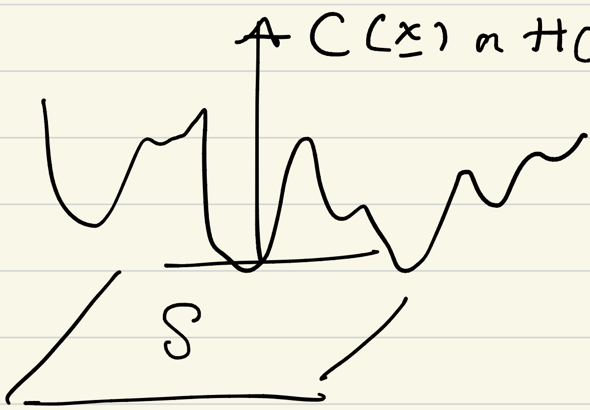
assignment $\underline{S} = (S_1, \dots, S_{|V|})$.

Cost $$H(\underline{S}) = -\sum_{(k,l)\in E} \underbrace{J_{kl}}_{\in \mathbb{R}} S_k S_l .$$

Goal find $\underline{S}^*$ s.t $H(\underline{S}^*) = \min_{\underline{S} \in \{\pm 1\}^V} H(\underline{S})$

$H : S = \{\pm 1\}^V \longrightarrow \mathbb{R}$.

$C(\underline{x})$ or $H(\underline{S}) \mapsto$ very complicated fcts in general with many many local minima & can have many global minima.

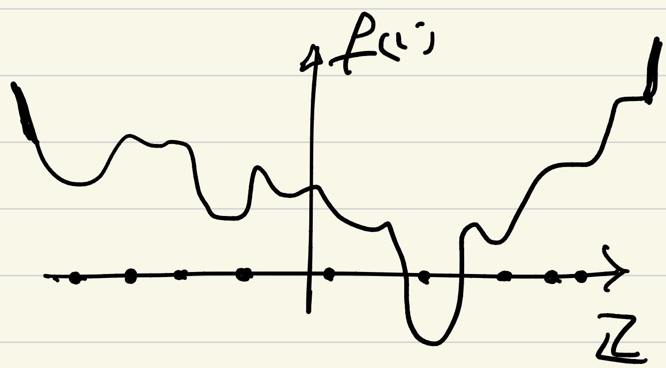


$\mapsto$ Algos (greedy) will typically be stuck in local minima.

. . . . . .

- To illustrate the concept of how to apply the Metropolis-Hasting algo & use an easy fct (one dimensional). But conceptually what comes next is general.

$$f : \mathbb{Z} \longrightarrow \mathbb{R}.$$
$$i \longmapsto f(i)$$



Assume $\lim_{i \to \pm\infty} f(i) = +\infty$.

"<u>First Idea</u>": use <u>sampling</u> of the uniform distr over global minima $\rightarrow$ By using some form of MCMC.

$$\pi_\infty(i) = \frac{\mathbb{1}\{i \text{ is a global min of } f\}}{Z_\infty}$$

$$Z_\infty = \sum_{i \in \underset{\substack{\uparrow \\ S}}{\mathbb{Z}}} \mathbb{1}\{i \text{ is a global min of } f\}.$$

Better Idea consider sampling by $\boxed{\text{MCMC}}$ of the following distribution

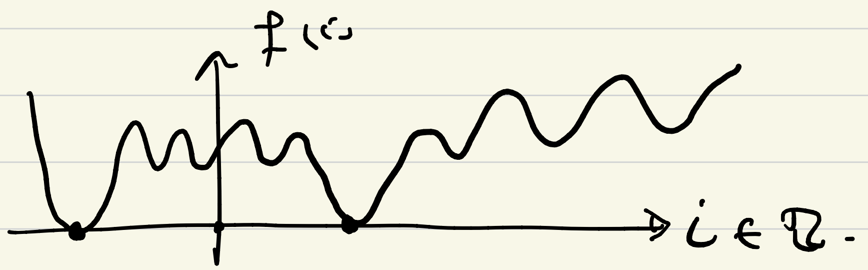
$$\pi_\beta(i) = \frac{e^{-\beta f(i)}}{Z_\beta} \quad ; \quad \beta > 0.$$

↓ interpret this later!

and $$Z_\beta = \sum_{i \in \mathcal{Z}} e^{-\beta f(i)}.$$

$\uparrow$ S

Remark: For $\beta \to +\infty$ $\quad \pi_\beta(i) \to \pi_\infty(i)$.

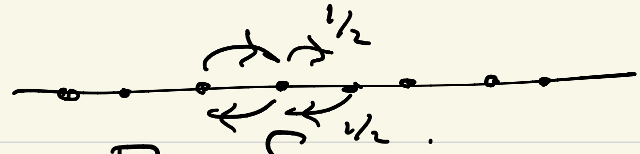


Not a loss of generality to set the problem s.t.

$$\min_{i \in \mathcal{Z}} f(i) = 0.$$

Remark: in fact if you think of Ising Model then clearly you can interpret β as "inverse temperature" and $f(i)$ as the "energy of state i" or "cost of state i".

Metropolis algorithm:



on $\mathbb{Z} = S$

1. Choose a base chain ψ. Here I propose to choose s.t. $\psi_{i,i\pm1} = \frac{1}{2}$ and zero otherwise.

   (sym random walk on $\mathbb{Z}$).

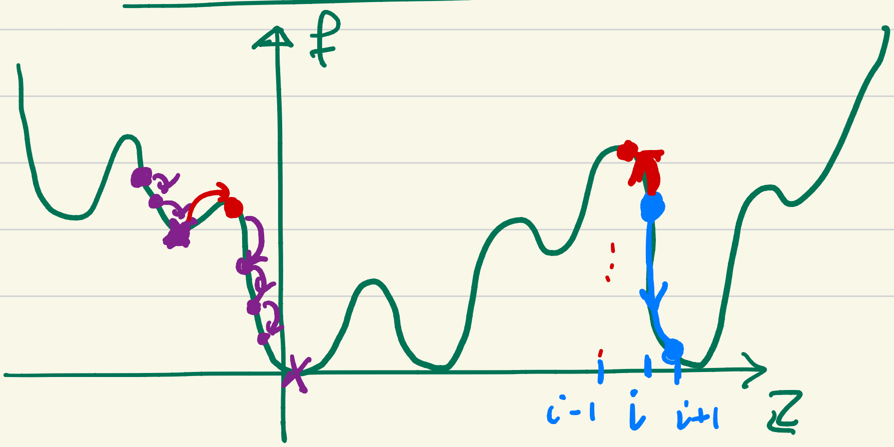
2. Accept moves with acceptance probabilities

$$a_{ij} = \min\left(1, \frac{\pi_j}{\pi_i}\right) = \begin{cases} \min\left(1, e^{-\beta(f(j)-f(i))}\right) & \text{if } j = i\pm1 \\ 0 & \text{otherwise.} \end{cases}$$

3. The following is the $\overbrace{\text{}}^{\text{new}}$ Markov Chain constructed

$$P_{ij} = \begin{cases} \psi_{ij}\, a_{ij} & i \neq j \\ 1 - \sum_{k \neq i} \psi_{ik} a_{ik}. & i = j \end{cases}$$

PICTURE OF WHAT HAPPENS:



- $f(i+1) - f(i) < 0 \Rightarrow a_{i\,i+1} = 1$ certainly accept this move.
- $f(i-1) - f(i) > 0 \Rightarrow a_{i\,i-1} < 1$ accept with prob <1 or you stay at i (self loop) with $1-\text{prob} < 1$.

Remarks:

- if a $\overbrace{\text{move}}^{\text{proposed}}$ lowers the cost then you certainly accept it.

- if a proposed move increases the cost, well you should sometimes accept it! That is the only way to get out of local minima and "un-stuck" yourself!

- What effect of β? "The inverse temperature"

✓ * If β is large then $a_{ij} = \min\left(1, e^{-\beta(f(j)-f(i))}\right)$

very low probability to accept a move that increases the cost. ← is very small if $f(j) > f(i)$
is equal to 1 if $f(j) < f(i)$

you don't have too many selfloops ~~→ chain is not~~
you have tendency of being stuck in local mins.

✓ * If β is small then $a_{ij} \simeq 1$. so you tend to always accept a move and chain will explore the whole state space and will tend π_β but for β small and this very different than $\underline{\underline{\pi_\infty}}$.

⇒ One must be careful how you choose β.

This is a kind of art. In the notes find

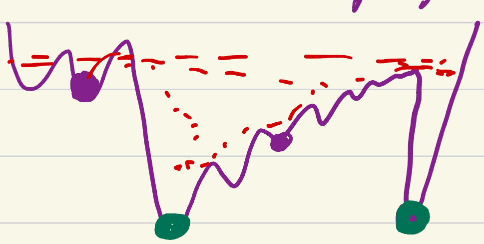
ballpark estimate how to choose it ✓

#.

<u>Simulated Annealing</u>: combine both β small and β large.

- Start with β_0 small (appropriate) you explore well the whole state space and after a reasonable amount of time you have almost converged to π_{β_0}.

$\pi_{\beta_0} = \dfrac{e^{-\beta_0 f(j)}}{Z_{\beta_0}} \approx$ unif distr over whole state space if $\beta = 0$.



<u>$\beta_0 \nearrow \beta_1$.</u>

- Then you increase β by a little. And run the new chain for some time. Converge approx π_{β_1}.
- Repeat. <u>$\beta_1 \nearrow \beta_2$</u>. Run the chain and approx conv to π_{β_2}.
- Repeat ---

$\Downarrow$ Eventually you reach samples distributed as almost π_∞ or $\pi_{\beta \text{ very large}}$ ⇒ These samples are low lying Minima. Maybe you can reach global Min.

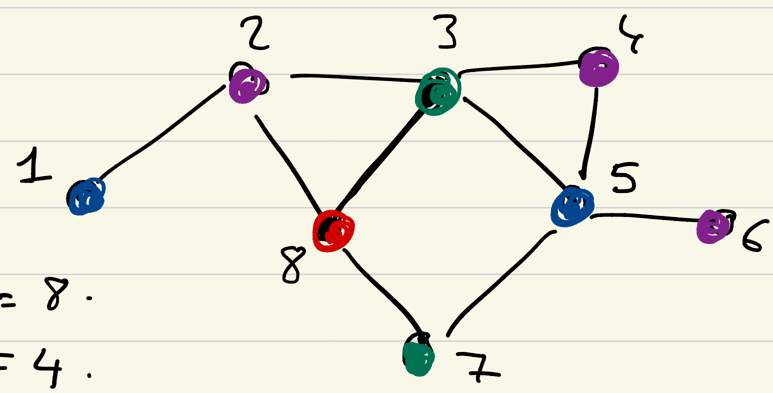
Introduction to graph coloring and MCMC.

Today : introduction & statement of a theorem.

Next time : Proof and analysis.

Recall the problem :

$G = (V, E)$



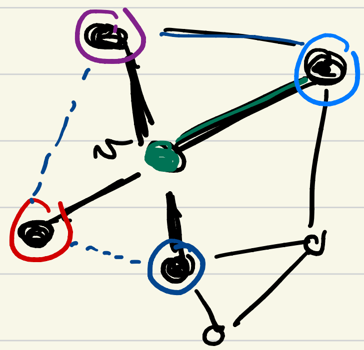
$\begin{cases} |V| = 8. \\ q = 4. \end{cases}$

- assignement of colors $\{1, 2, 3, \dots q\}$ is a vector $(x_1, x_2, \dots x_{|V|}) = \underline{x}$ where and $x_k \in \{1, 2, 3, \dots q\}$.
- Proper coloring a proper assignement is s.t any $(k, l) \in E$ are s.t $x_k \neq x_l$.
- Aim is to sample from :

$$\pi(\underline{x}) = \frac{\mathbb{1}(\underline{x} \text{ is a proper } q\text{-coloring})}{Z}.$$ ✓

Remark: arbitrary $G = (V, E)$ and arbitrary q it is not always the case that a proper col exists.

But a least if $q \geq \max_{v \in V} \deg(v) + 1$ then a proper coloring exist.



$\deg(v) = 4$.

Notation: $\max_{v \in V} \deg(v) \equiv \Delta$.

For now on $q \geq \Delta + 1$.

Remark: Space of proper q-coloring is called

$$S = \{ \underline{x} = (x_1, \dots x_{|V|}) \mid \forall (k, l) \in E \quad x_k \neq x_l \}.$$

↑ will play the role of the state-space of the MC.

[$S \subset$ space of all assignements.]

- We want to sample uniformly from S.
- The following Algorithm is used and we will prove that indeed this works at least for q large enough.
- $\boxed{\text{Algorithm.}}$

1. Start from a proper q-col $\underline{x} \in S$. $(*)$.
2. Select a vertex $v \in V$ uniformly at random.
3. Select a color $c \in \{1, \dots, q\}$ unif at random.
4. If c is allowed at v then recolor vertex v (that is $\underset{\text{old color}}{x_v} \to \underset{\text{New color.}}{c}$). If c is not allowed then you do nothing.

- We will convince ourselves later that this is an instance of a Metropolis Algo.

- First step is not really a problem because one can always start from any assignment $\underline{x}$ and proceed as the algo tells you; apply steps 2 → 4.
- Nb of edges that have a wrong color can only decrease, or stay equal.
- Eventually you will converge to $\mathcal{S}$.

#.

Theorem. If $q > 3\Delta$, then for all proper (initial) colorings $\underline{x}$:

$$\| P_{\underline{x}}^{m} - \pi \|_{TV} \leq |V| e^{-\frac{m}{|V|}\left(1 - \frac{3\Delta}{q}\right)}$$

prob distr of chain starting at $\underline{x}$ running for m-steps

unif prob distr over proper col.

Furthermore $$T_\epsilon \leq \frac{1}{1 - \frac{3\Delta}{q}} |V| \left(\log |V| + \log \frac{1}{\epsilon} \right).$$

Remarks: optimal q_c s.t algo works given Δ.

is it $\Delta+1$?

- In general this question is difficult and depends on G.
- Thm works for all G's.
- Interesting classes of graphs

  e.g., Erdos-Rényi graph, $q_c(\Delta)$

  can be determined.

* We limit ourselves to $\underbrace{\text{arbitrary graphs}}_{q > 3\Delta}$.

* Book Levin, Peres you find a better condition and

  a better thm. However analysis is more difficult.

Remark: We prove the thm (next time) with no reference to general theory of ergodic MC's.

Last thing for today: recognize this also as as Metropolis algo.

1. Start from $\underline{x}$ proper q-col.
2. Select $v \in V$ at random.
3. Select $c \in \{1, \dots, q\}$ at random
4. If c is allowed you recolor v. Otherwise do nothing

Repeat.

Base chain?

$$\underline{x} \bullet \longrightarrow \bullet \, \underline{y} . \qquad \underline{x} \bullet \circlearrowleft$$

$$\psi_{\underline{x}\,\underline{y}} = \frac{1}{|V|} \frac{1}{q}$$; if $\underline{y}$ is such that $\underline{x}$ and $\underline{y}$ differ at only one vertex, $\underline{\underline{v}}$,

Neighboring cols in S.

$$\psi_{\underline{x}\underline{x}} = 1 - \sum_{\underline{y} \neq \underline{x}} \psi_{\underline{x}\,\underline{y}}$$ if $\underline{x} = \underline{y}$ self loop.

$$= 1 - \frac{1}{q|V|} \underbrace{\sum_{\underline{y} \neq \underline{x}} \psi_{\underline{x}\,\underline{y}}}_{|V|(q-1)}$$

$$= 1 - \frac{|V|(q-1)}{q|V|} = 1 - \left(1 - \frac{1}{q}\right) = \frac{1}{q}.$$

$\psi_{\underline{x}\,\underline{y}} = 0$ If $\underline{y}$ & $\underline{x}$ are not Neighbored or equal.

acceptance prob: $\quad (\psi_{\underline{x}\underline{y}} = \psi_{\underline{y}\underline{x}})$

$$a_{\underline{x}\underline{y}} = \min\left(1; \frac{\pi_{\underline{y}}}{\pi_{\underline{x}}}\right)$$

$$= \min\left(1; \frac{\mathbb{1}(\underline{y} \text{ is a proper col})}{\cancel{Z}} \cdot \frac{1}{\frac{\mathbb{1}(\underline{x} \text{ is proper})}{\cancel{Z}}}\right)$$

$= 1$ at each step of chain

$$\Rightarrow a_{\underline{x}\underline{y}} = \min\left(1; \mathbb{1}(\underline{y} \text{ is a proper col})\right)$$

$$= \mathbb{1}(\underline{y} \text{ is a proper col}).$$

$\downarrow$

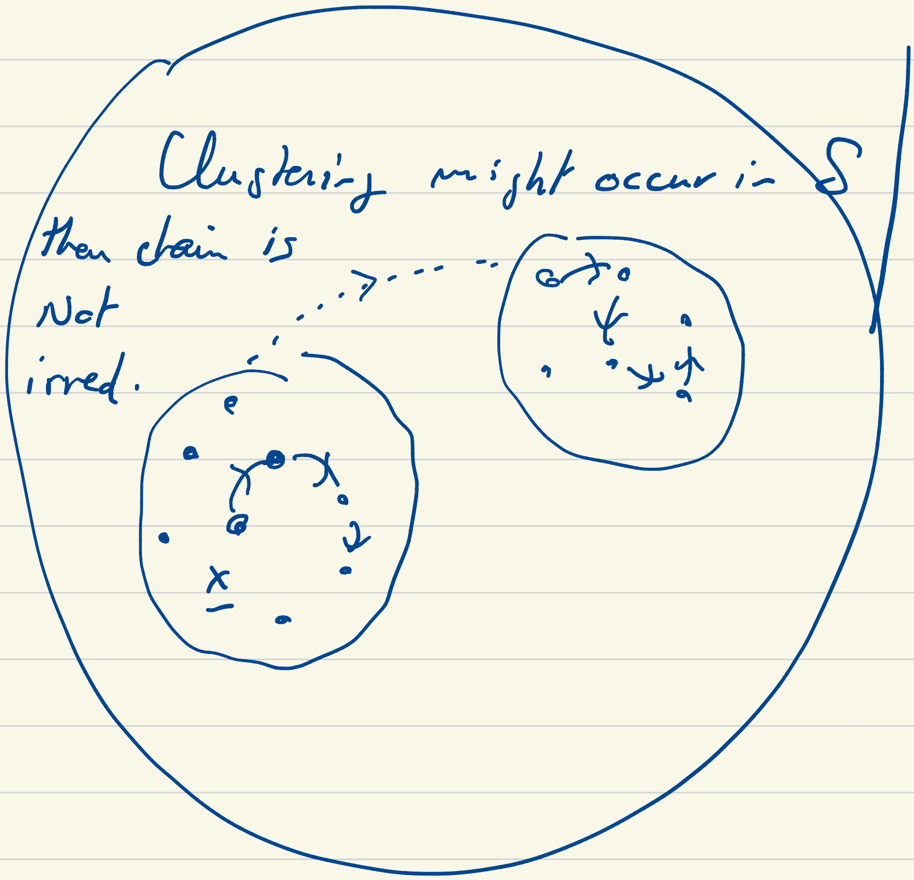
This means that if the recolor of v is allowed $a_{\underline{x}\underline{y}} = 1$

and if the recolor of v is not allowed $a_{\underline{x}\underline{y}} = 0$ and do nothing.

$\blacksquare$

Remark: Not easy to see if the chain is irreducible. Indeed we will not have to prove this because we prove theorem by "bare hands".

But if $q > 3\Delta \to$ irred:



S space of all proper col.

∎