# INFERENCE USING MARKOV CHAIN MONTE CARLO

For linear regression, we derived the least squares formula by maximizing

$p(\underline{y} | \underline{w}, \beta)$ i.e. probability of getting some output given a model

to infer the best parameters $\underline{w}$ given some data we can try to learn the POSTERIOR distribution $p(\underline{w} | y)$ (we neglect $\beta$ for now)

which can be written as

$$p(\underline{w}|\underline{y}) = p(\underline{w})\, p(\underline{y}|\underline{w})$$

and taking the log

$$\ln\left[p(\underline{w}|\underline{y})\right] = \underbrace{\ln\left[p(\underline{w})\right]}_{\text{prior probability of weights}} + \underbrace{\ln\left[p(\underline{y}|\underline{w})\right]}_{\text{least squares formula}}$$

$p(\underline{w})$ can be chosen using some ansatz e.g. uniform, gaussian, ...

the idea is to sample the parameter space $p(\underline{w})$ to look for a minimum of $\ln(p(\underline{w}|\underline{y}))$ which provides a best fit parameter choice

The sampling can be done using <u>Markov chain Monte Carlo</u> methods where each point $x(t_i) = \{w_i\}$ in the chain depends only on the position of the previous one