Score Fn

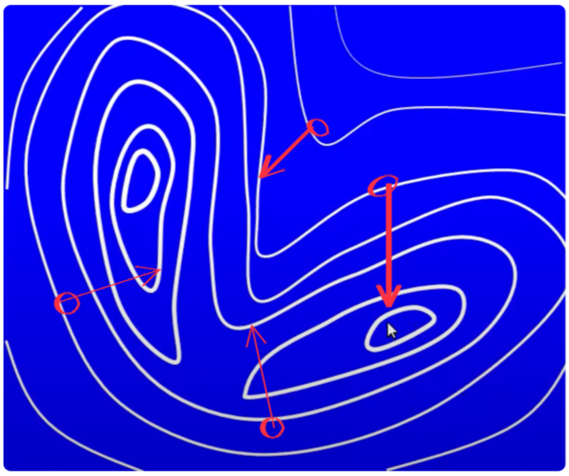
Hello to Score function

$$P(x)$$

Score function $\nabla_x \log P(x)$ $\longleftrightarrow$ $\hat{x} = G_\theta(z)$

gradient of $P(x)$

This function is highly informative, as it points to the dir^n of best ascent on function $\log P(x)$ and thus also on $P(x)$.



If this is the distribution $P(x)$ then,
the $\nabla_x \log P(x)$ points up the hill in fastest possible way.

How relevant is the Score function?

If we had access to the Score function, we could solve (MAP-wise) various inverse problem in straight forward manner, and without a need of prior.

$$\hat{x}_{k+1} = \hat{x}_k + \mu\left[\frac{1}{\sigma^2} C^T(y - C\hat{x}_k) - \boxed{\nabla_x \log P(x)\Big|_{x_k}}\right]$$

$$\boxed{\hat{x}_k = \underset{x}{\mathrm{Argmin}}\ \frac{1}{2\sigma^2}\|y - Cx\|_2^2 - \log P(x)}$$

☆ Also, in Gibb's form $\left(P(x) = \frac{1}{Z}\exp[-\varphi(x)]\right)$

$$\boxed{\nabla_x \log P(x) = -\nabla_x f(x)}$$

Why is this exciting? Because the score function skips the need to calculate the partition function (Z), which is a source of difficulties

⇒ After some derivation,

$$\sigma^2 \nabla_y \log P(y) = \int_x (x-y)\, P(x|y)\, dx$$

opening this,

$$\sigma^2 \nabla_y \log P(y) = \sigma^2 \frac{\nabla_y P(y)}{P(y)} = \int_x x P(x|y)\, dx - y \int_x P(x|y)\, dx$$

∴ $E(x|y)$ is the MMSE denoiser

and this is Conditional Expectation

is equal to 1 for any given y, x will take on some value with 100% certainty.

So, $\boxed{\sigma^2 \nabla_y \log P(y) = E(x|y) - y}$

→ getting the score is simple - Apply your (MMSE) denoiser on y and subtract this image.
i.e. provide the estimated noise from the denoiser.

☆ MMSE denoiser is get from training a denoiser over the images.

Overall → with a pre trained such denoiser, we can approximate the score function quite well.