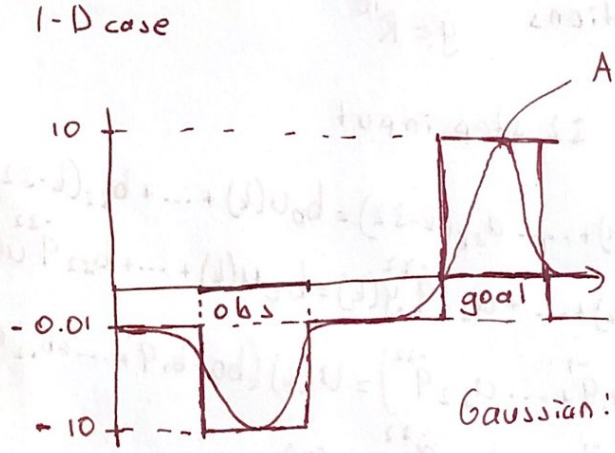
Reward System

1-D case



Gaussian: $\frac{1}{\sqrt{2\pi}\,|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$ (Σ: 2×2, $(x-\mu)$: 2×1, $(x-\mu)^T$: 1×2)

$\frac{1}{\sqrt{(2\pi)^N}\,|\Sigma|^{1/2}} e^{-\frac{1}{2}x^T \Sigma^{-1} x}$

$r: R^2 \to R$

Σ = identity

$$r = \underbrace{-0.01}_{\text{step cost}} + \underbrace{A_o e^{-B_o (x-\mu_o)^T (x-\mu_o)}}_{\text{collision penalty for each obs}} + \underbrace{A_g e^{-B_g (x-\mu_g)^T (x-\mu_g)}}_{\text{goal reward}}$$

at goal $x = \mu_g$

$A_g = 20$

B_g

at obs $x = \mu_o$

$A_o = -20$

B_o

$\Rightarrow \Sigma = \sigma^2 I$

$\Sigma^{-1} = \frac{1}{\sigma^2} I$

$B = \frac{1}{2\sigma^2} = \frac{2}{rad^2}$

radius

$r = \begin{cases} 20, & \text{goal} \\ -20, & \text{collision} \\ -0.01, & \text{else} \end{cases}$

$radius = 2\sigma$

$\sigma = \frac{rad}{2}$

=> Borders

End simulation with no penalty?