$$\pi(a|s,\theta) = \frac{1}{\sigma(s,\theta)\sqrt{2\pi}} \exp\left(-\frac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right)$$

Eq 13.19 from RL book.

Here: $\mu(s,\theta) = \theta_\mu^T x_\mu(s)$

$\sigma(s,\theta) = \exp(\theta_\sigma^T x_\sigma(s))$ → Eq 13.20 from RL book

To prove:-

$$\nabla \ln \pi(a|s,\theta_\mu) = \frac{\nabla \pi(a|s,\theta_\mu)}{\pi(a|s,\theta)} = \frac{1}{\sigma(s,\theta)^2}(a-\mu(s,\theta))x_\mu(s) \quad \text{(i)}$$

Taking L.H.S from eq i)

$$\frac{\nabla \pi(a|s,\theta_\mu)}{\pi(a|s,\theta)} = \frac{1}{\pi(a|s,\theta)} \frac{\partial \pi(a|s,\theta_\mu)}{\partial \theta_\mu}$$

Using identity $\frac{d}{dx} U \cdot V = U\frac{dV}{dx} + V\frac{dU}{dx}$

$$= \frac{1}{\pi(a|s,\theta)} \cdot \frac{1}{\sigma(s,\theta)\sqrt{2\pi}} \frac{d}{d\theta_\mu} \exp\left(-\frac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) + 0$$

$$= \frac{1}{\pi(a|s,\theta)} \cdot \frac{1}{\sqrt{2\pi}\,\sigma(s,\theta)} \exp\left(-\frac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) \left[\frac{2\sigma(s,\theta)^2[-2(a-\mu(s,\theta))(-x_\mu(s))]}{[2\sigma(s,\theta)^2]^2}\right]$$

$$= \frac{1}{\frac{\cancel{\exp\left(-\frac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right)}}{\cancel{\sigma(s,\theta)\sqrt{2\pi}}}} \times \frac{1}{\cancel{\sqrt{2\pi}\,\sigma(s,\theta)}} \cancel{\exp\left(-\frac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right)} \left[\frac{\cancel{2 \cdot 2\sigma(s,\theta)^2}(a-\mu(s,\theta))\,x_\mu(s)}{2\sigma(s,\theta)^2 \cancel{2\sigma(s,\theta)^2}}\right]$$

$$= \frac{(a-\mu(s,\theta))\,x_\mu(s)}{\sigma(s,\theta)^2} \quad \square$$

To prove :-

$$\nabla \ln \pi(a|s,\theta_{\sigma}) = \frac{\nabla \pi(a|s,\theta_{\sigma})}{\pi(a|s,\theta_{\sigma})} = \left(\frac{(a-\mu(s,\theta))^2}{\sigma(s,\theta)^2} - 1\right) x_{\sigma}(s) \quad \text{ii}$$

Taking L.H.S from eq ii :-

$$\frac{\nabla \pi(a|s,\theta_{\sigma})}{\pi(a|s,\theta_{\sigma})} = \frac{1}{\pi(a|s,\theta_{\sigma})} \frac{\partial}{\partial \theta_{\sigma}} \pi(a|s,\theta_{\sigma})$$

using identity $\frac{d}{dx} U \cdot V = U\frac{dV}{dx} + V\frac{dU}{dx}$

$$\frac{d}{dx}\frac{U}{V} = \frac{V\frac{dU}{dx} - U\frac{dV}{dx}}{V^2}$$

$$= \frac{1}{\pi(a|s,\theta_{\sigma})}\left[\frac{1}{\sigma(s,\theta)\sqrt{2\pi}} \frac{d}{d\theta_{\sigma}} \exp\left(\frac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) + \exp\left(\frac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) \frac{d}{d\theta_{\sigma}} \frac{1}{\sigma(\theta,s)\sqrt{2\pi}}\right]$$

$$= \frac{1}{\pi(a|s,\theta_{\sigma})}\left[\frac{1}{\sigma(s,\theta)\sqrt{2\pi}}\left\{\frac{0+(a-\mu(s,\theta))^2 \, 2 \cdot 2\sigma(s,\theta)\sigma(s,\theta)\, x_{\sigma}(s)}{4(\sigma(s,\theta)^2)^2}\right\}\right.$$

$$\left. + \exp\left(\frac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right)\left\{\frac{\sigma(s,\theta)\sqrt{2\pi}\times 0 - 1\sqrt{2\pi}\exp(\theta_{\sigma}^T x_{\sigma}(s))\, x_{\sigma}(s)}{(\sigma(s,\theta)\sqrt{2\pi})^2}\right\}\right]$$

$$= \frac{1}{\pi(a|s,\theta_{\sigma})}\left[\frac{1}{\sigma(s,\theta)\sqrt{2\pi}}\left\{\frac{\exp\left(\frac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) 4\sigma(s,\theta)(a-\mu(s,\theta))^2 \exp(\theta_{\sigma}^T x_{\sigma}(s))\, x_{\sigma}(s)}{4(\sigma(s,\theta)^2)^2}\right\}\right.$$

$$\left. + \exp\left(\frac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right) \frac{-\sqrt{2\pi}\, x_{\sigma}(s)\exp(\theta_{\sigma}^T x_{\sigma}(s))}{(\sigma(s,\theta)\sqrt{2\pi})^2}\right]$$

$$= \frac{\dfrac{1}{\pi(a|s,\theta)} \times \dfrac{(a-\mu(s,\theta))^2 x_\sigma(s)}{\sigma(s,\theta)^3\sqrt{2\pi}} \cancel{\exp\left(\dfrac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)}\right)} + \cancel{\exp\left(\dfrac{-(a-\mu(s,\theta))^2}{2\sigma(s,\theta)^2}\right)} \times \dfrac{-x_\sigma(s)}{\sigma(s,\theta)\sqrt{2\pi}}}{\dfrac{\cancel{\exp\left(-\dfrac{(a-\mu(s,\theta))^2}{2\sigma(s,\theta)}\right)}}{\sigma(s,\theta)\sqrt{2\pi}}}$$

$$= \frac{(a-\mu(s,\theta))^2 x_\sigma(s)}{\sigma(s,\theta)^2} + - x_\sigma(s)$$

$$= \left(\frac{(a-\mu(s,\theta))^2}{\sigma(s,\theta)^2} - 1\right) x_\sigma(s) \qquad \square$$