

# Chapter 1

## Derivations

### 1.1 Natural policy gradient

The TRPO objective is

$$\max_{\theta} \mathbb{E}_{s_0, \dots, s_{H-1} \sim \rho_{\theta k}}$$

finish derivat  
from 2023-1  
lecture