Deriving Optimal Discriminator

Derivation Steps

1 Fix generator G, optimize D

$$\max_D V(D,G)$$

↓

2 Calculus of variations

$$\partial V/\partial D(x) = 0$$

↓

3 Solve for optimal D

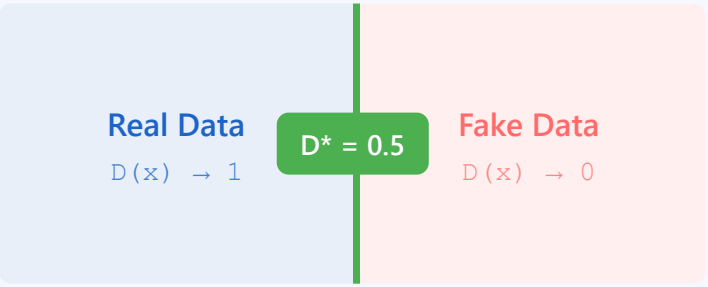
$$p_{data}(x)/D(x) = p_g(x)/(1-D(x))$$

✓ Optimal Discriminator

Key Insights

- Sigmoid activation natural choice
- Binary cross-entropy loss
- Depends on density ratio
- Neural network approximation

Decision Boundary



At equilibrium: $p_g = p_{data}$

$$D^*(x) = p_{data}(x) / (p_{data}(x) + p_g(x))$$