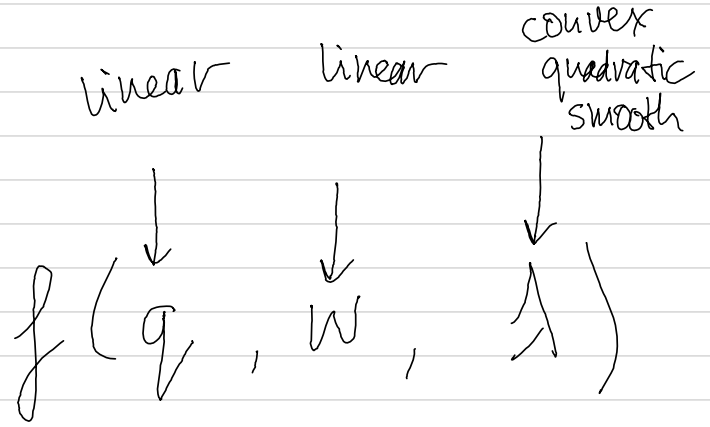
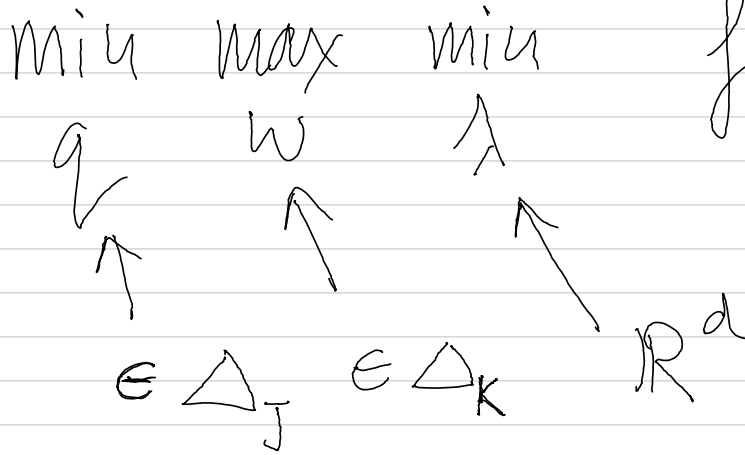


Wouter Koolen



Optimization for pure exploration



# Best Arm Identification

seq. dec. problem.

(Drug testing, A/B testing, )

$K$  "arms"

hidden vector  $\mu \in \mathbb{R}^K$  mean rewards, Gaussian  $\sigma=1$

Interaction protocol:

for  $t=1, 2, \dots, \tau$

pick  $A_t \in [K]$

see  $X_t \stackrel{\text{indep}}{\sim} \mathcal{N}(\mu_{A_t}, 1)$

output recommendation  $\hat{c} \in [K]$

→ confidence ← fix

→ fast ← optimize

look at  $\delta$ -correct algorithms

$$\forall \mu : \mathbb{P}_\mu (\hat{c} \neq i^*(\mu)) \leq \delta$$

↓  
argmax<sub>k</sub>  $\mu_k$

$$(A_1, X_1, \dots, A_\tau, X_\tau, \hat{c}) =: H$$

the sample complexity at  $\mu$  is  $\mathbb{E}_\mu[\tau]$

lower bds either at  $\mu$  or at  $\lambda$   $i^*(\mu) \neq i^*(\lambda)$

$$\begin{aligned} & \text{KL} \left( \mathbb{P}_\mu^{\text{alg}}(H) \parallel \mathbb{P}_\lambda^{\text{alg}}(H) \right) \\ & \geq \underbrace{\text{KL} \left( \mathbb{P}_\mu^{\text{alg}}(\hat{c} = i^*(\mu)) \parallel \mathbb{P}_\lambda^{\text{alg}}(\hat{c} = i^*(\mu)) \right)}_{\geq 1-\delta} \underbrace{\left( \mathbb{P}_\lambda^{\text{alg}}(\hat{c} = i^*(\mu)) \right)}_{\leq \delta} \\ & \geq \\ & \approx \ln 1/\delta \end{aligned}$$

$$\begin{aligned} & = \sum_{a=1}^K \underbrace{\mathbb{E}_\mu[\mathbb{N}_a(\tau)]}_{\text{does not depend on } \lambda} \underbrace{\text{KL}(\mu_a \parallel \lambda_a)}_{\frac{1}{2}(\mu_a - \lambda_a)^2} \end{aligned}$$

so:  $\inf_{\lambda: i^*(\lambda) \neq i^*(\mu)} \geq \ln 1/\delta$

$$\mathbb{E}_\mu[\tau] \geq \inf_{\lambda \dots} \sum_{a=1}^K \frac{\mathbb{E}[\mathbb{N}_a(\tau)]}{\mathbb{E}[\tau]} \text{KL}(\mu_a, \lambda_a) \geq \ln 1/\delta$$

$$\mathbb{E}_\mu[\tau] \geq \sup_{\text{WEAK } \lambda \dots} \inf \left\{ \sum_{a=1}^K w_a \text{KL}(\mu_a, \lambda_a) \right\} \geq \ln 1/\delta$$

$$\mathbb{E}_\mu[\tau] \geq \frac{\ln 1/\delta}{f_a(\mu)}$$