

Minimax Algorithm for Learning Rotations

Wojciech Kotłowski

CWI

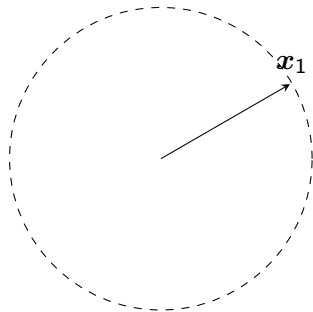
Manfred K. Warmuth

UCSC

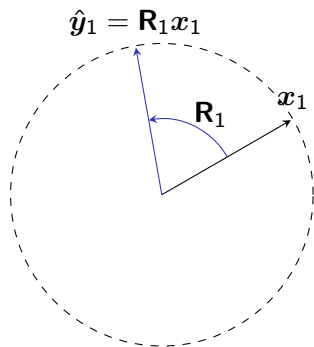
Open problem at COLT 2011

July 9, 2011

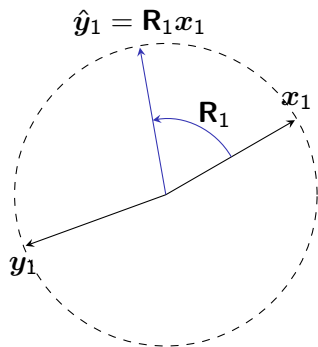
On-line Protocol



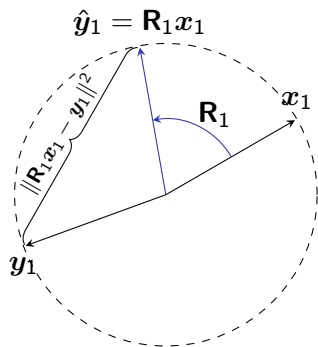
On-line Protocol



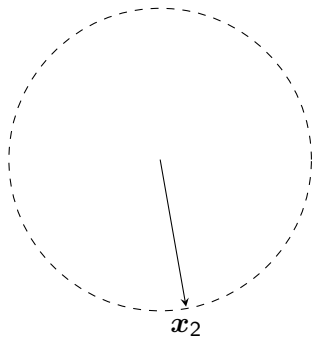
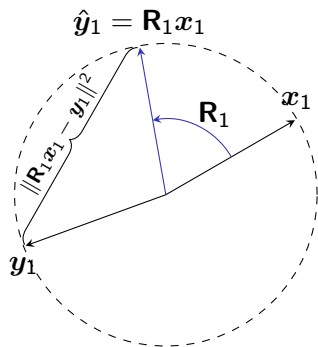
On-line Protocol



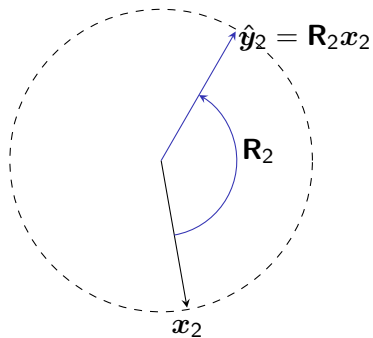
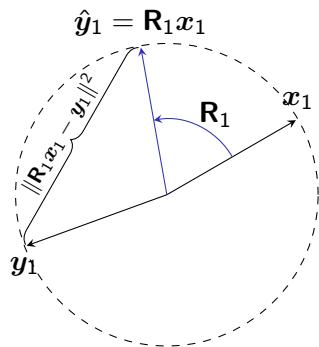
On-line Protocol



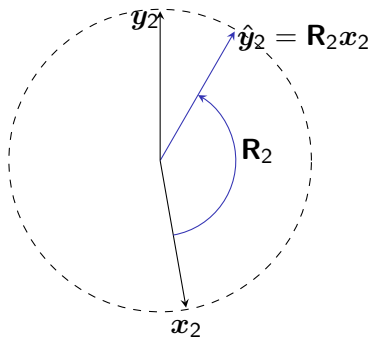
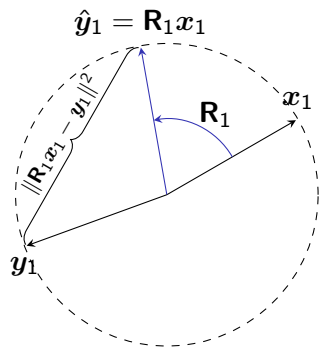
On-line Protocol



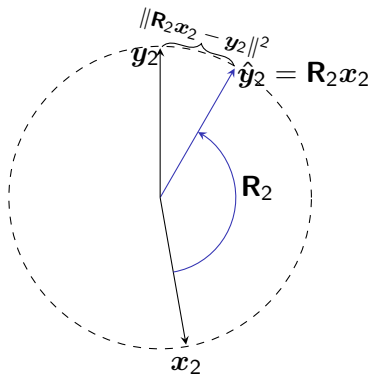
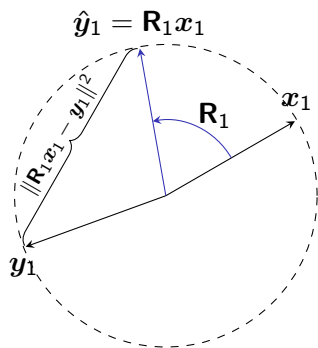
On-line Protocol



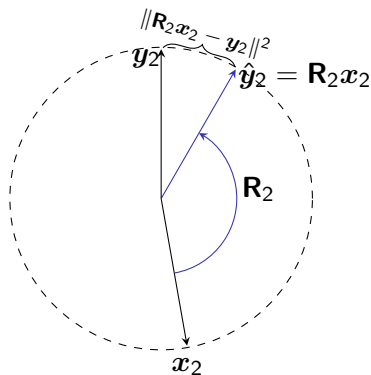
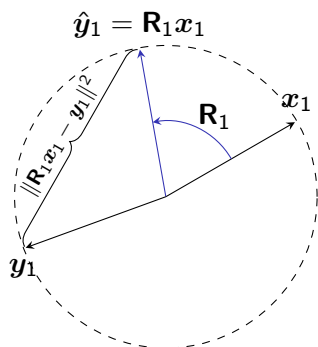
On-line Protocol



On-line Protocol



On-line Protocol



...

Goal: minimize regret

$$\sum_{t=1}^T L_t(\mathbf{R}_t) - \min_{\mathbf{R} \in \mathcal{SO}(n)} \sum_{t=1}^T L_t(\mathbf{R})$$

Applications in vision, robotics, matrix completion, subspace tracking, quantum physics, etc. [Arora 2009, Hazan et al. 2010]

Problem Setting

- Set of rotation matrices $\mathcal{SO}(n)$ is *non-convex*, but is a *Lie group* equipped with *Lie algebra*
- Loss is linear:

$$\begin{aligned}L_t(\mathbf{R}_t) &= \frac{1}{2} \mathbb{E} [\|\mathbf{R}_t \mathbf{x}_t - \mathbf{y}_t\|^2] \\ &= \frac{1}{2} \mathbb{E} [\underbrace{\|\mathbf{R}_t \mathbf{x}_t\|^2}_1 + \underbrace{\|\mathbf{y}_t\|^2}_1 - 2(\mathbf{R}_t \mathbf{x}_t) \cdot \mathbf{y}_t] \\ &= 1 - (\mathbb{E}[\mathbf{R}_t] \mathbf{x}_t) \cdot \mathbf{y}_t\end{aligned}$$

- Batch problem is known as *Wahba's problem*
Solvable with SVD

Previous Work

- [Arora, NIPS '09]: *Matrix exponentiated gradient algorithm*

- exploits Lie group/algebra structure
- deterministic
- problems with the motivation

[Hazan et al., COLT '10]:

- Any deterministic algorithm \Rightarrow can be forced to have regret $\Omega(T)$

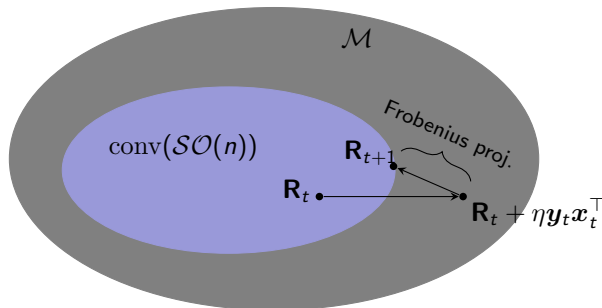
Previous Work

- [Arora, NIPS '09]: *Matrix exponentiated gradient algorithm*

- exploits Lie group/algebra structure
- deterministic
- problems with the motivation

[Hazan et al., COLT '10]:

- Any deterministic algorithm \Rightarrow can be forced to have regret $\Omega(T)$
- Randomized gradient descent:
 - Frobenius norm regularization with projection
 - $\sqrt{2nT}$ regret



Problem Statement

Is the problem already solved?

- Is Frobenious norm proper way to regularized rotations?
- Is there a notion of entropy defined over $\mathcal{SO}(n)$?
- How to maintain uncertainty over rotation matrices?
- Elegant Lie group and Lie algebra structure not exploited

We hope to shed some light on these issues by finding the minimax algorithm for learning rotations

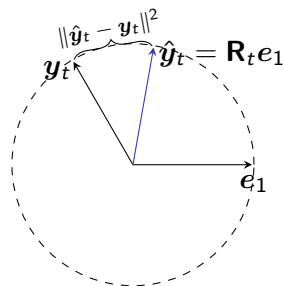
Partial Results

- If the instances x_t are restricted to be a fixed unit vector, say e_1 , then we can give the minimax algorithm:
 - Enough to predict with \hat{y}_t , s.t. $\|\hat{y}_t\| \leq 1$
 - The minimax prediction:

$$\hat{y}_t = \frac{s_{t-1}}{\sqrt{\|s_{t-1}\|^2 + T - t + 1}},$$

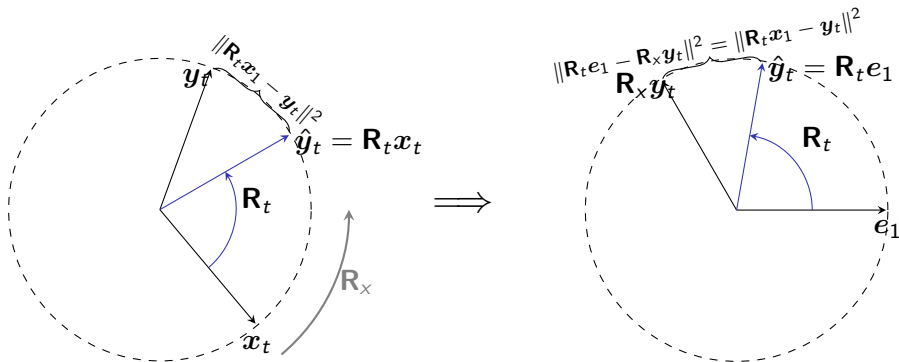
$$\text{where } s_{t-1} = \sum_{q=1}^{t-1} y_q$$

Minimax regret: \sqrt{T} .



Partial Results

For $n = 2$, a simpler problem is equivalent to learning rotations.



$$\|R_t e_1 - R_x y_t\|^2 = \|R_t R_x x_t - R_x y_t\|^2 \stackrel{(n=2)}{=} \|R_x R_t x_t - R_x y_t\|^2 = \|R_t x_t - y_t\|^2$$

For $n > 2$, the transformation does not work anymore.

Minimax algorithm
for learning rotations
when $n \geq 2$