# 1.1 Introduction

# 1.2 Définition formelle de l’estimateur NCE

# 1.3 Link with logistic regression

1.4 Some properties of NCE

The weak law of large numbers shows that as Td increases the objective function JT (θ) converges in probability to J, J(θ) = E{ln[h(x;θ)]}+ν E{ln[1−h(y;θ)]}.

Theorem 2 (Consistency) If conditions (a) to (c) are fulfilled then θˆ T converges in probability to θ ⋆ , θˆ T P → θ ⋆

Corollary 5 For ν → ∞, Σ is independent of the choice of pn and equals Σ = I −1 −I −1E(g)E(g) TI −1 ,

Corollary 7 If pn = pd then Σ = 1 + 1 ν I −1 −I −1E(g)E(g) TI −1 .

# 1.5 Implementation of NCE

- One dimensional case , table showing results of estimation , stochastic gradient descent

- Multidimensional case , same thing, automatic differentiation with Tensorflow

# 1.6 Numerical experiments and simulation

## 1.6.1 Impact of nu

1.6.2 Choice of noise

Graph with KL distance

Discussion on the importance of the choice of noice

Transition with GAN