"Broadening the knowledge base and supporting the long term professional sustainability of the Research University Centre of Excellence at the University of Szeged by ensuring the rising generation of excellent scientists.

ÚJ SZÉCHENYI TERV

Doctoral School of Mathematics and Computer Science

Stochastic Days in Szeged

27.07.2012.

Online-to-Confidence-Set Conversions and Application to Sparse Stochastic Bandits

Csaba Szepesvári

(University of Alberta)

UNIVERSITAS SCIENTIARUM SZEGEDIENSIS · SZEGEDI TUDOMÁNYEGYETEM

MAGYARORSZÁG MEGÚJUL

A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

TÁMOP-4.2.2/B-10/1-2010-0012 project

Online-to-Confidence-Set Conversions and Application to Sparse Stochastic Bandits

Csaba Szepesvári

Department of Computing Science
University of Alberta
csaba.szepesvari@ualberta.ca
UNIVERSITY OF ALBERTA

July 27, 2013

Stochastic days – honoring András Krámli 70th birthday

joint work with Yasin Abbasi-Yadkori and Dávid Pál

Contents

- Linear prediction and (honest) confidence sets

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Linear Prediction and (Honest) Confidence Sets

Getting directions from András: MDPs, ILT, Dregely

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

- The "noise", η_t is conditionally R-sub-Gaussian with some $R > 0$, i.e.,

$$\forall \lambda \in \mathbb{R} : \qquad \mathbf{E}[e^{\lambda \eta_t} | X_1, \ldots, X_t, \eta_1, \ldots, \eta_{t-1}] \leq \exp\left(\frac{\lambda^2 R^2}{2}\right) .$$

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

- The "noise", η_t is conditionally R-sub-Gaussian with some $R > 0$, i.e.,

$$\forall \lambda \in \mathbb{R}: \qquad \mathbf{E}[e^{\lambda \eta_t} | X_1, \ldots, X_t, \eta_1, \ldots, \eta_{t-1}] \leq \exp\left(\frac{\lambda^2 R^2}{2}\right) .$$

- Often X_t is chosen based on $(X_1, \ldots, X_{t-1})$ and $(Y_1, \ldots, Y_{t-1})$

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

- The "noise", η_t is conditionally R-sub-Gaussian with some $R > 0$, i.e.,

$$\forall \lambda \in \mathbb{R} : \qquad \mathbf{E}[e^{\lambda \eta_t} | X_1, \ldots, X_t, \eta_1, \ldots, \eta_{t-1}] \leq \exp\left(\frac{\lambda^2 R^2}{2}\right) .$$

- Often X_t is chosen based on $(X_1, \ldots, X_{t-1})$ and $(Y_1, \ldots, Y_{t-1})$

Estimation Problems:

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

- The "noise", η_t is conditionally R-sub-Gaussian with some $R > 0$, i.e.,

$$\forall \lambda \in \mathbb{R}: \qquad \mathbf{E}[e^{\lambda \eta_t} | X_1, \ldots, X_t, \eta_1, \ldots, \eta_{t-1}] \leq \exp\left(\frac{\lambda^2 R^2}{2}\right) .$$

- Often X_t is chosen based on $(X_1, \ldots, X_{t-1})$ and $(Y_1, \ldots, Y_{t-1})$

Estimation Problems:

- Estimate θ_* based on $((X_1, Y_1), \ldots, (X_n, Y_n))$!

The Data

- $X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}$
- $\exists \theta_* \in \mathbb{R}^d$ s.t.

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t, \quad t = 1, \ldots, n$$

- The "noise", η_t is conditionally R-sub-Gaussian with some $R > 0$, i.e.,

$$\forall \lambda \in \mathbb{R} : \qquad \mathbf{E}[e^{\lambda \eta_t} | X_1, \ldots, X_t, \eta_1, \ldots, \eta_{t-1}] \leq \exp\left(\frac{\lambda^2 R^2}{2}\right) .$$

- Often X_t is chosen based on $(X_1, \ldots, X_{t-1})$ and $(Y_1, \ldots, Y_{t-1})$

Estimation Problems:

- Estimate θ_* based on $((X_1, Y_1), \ldots, (X_n, Y_n))$!
- Construct a confidence set that contains θ_* w.h.p.!

Sub-Gaussianity

Definition

Random variable Z is *R-sub-Gaussian* for some $R \geq 0$ if

$$\forall \gamma \in \mathbb{R} \qquad \mathbf{E}[e^{\gamma Z}] \leq \exp\left(\frac{\gamma^2 R^2}{2}\right) .$$

Sub-Gaussianity

Definition

Random variable Z is R*-sub-Gaussian* for some $R \geq 0$ if

$$\forall \gamma \in \mathbb{R} \qquad \mathbf{E}[e^{\gamma Z}] \leq \exp\left(\frac{\gamma^2 R^2}{2}\right) .$$

The condition implies that

- $\mathbf{E}[Z] = 0$
- $\mathbf{Var}[Z] \leq R^2$

Sub-Gaussianity

Definition

Random variable Z is R*-sub-Gaussian* for some $R \geq 0$ if

$$\forall \gamma \in \mathbb{R} \qquad \mathbf{E}[e^{\gamma Z}] \leq \exp\left(\frac{\gamma^2 R^2}{2}\right) .$$

The condition implies that

- $\mathbf{E}[Z] = 0$
- $\mathbf{Var}[Z] \leq R^2$

Examples:

- Zero-mean bounded in an interval of length $2R$ (Hoeffding-Azuma)
- Zero-mean Gaussian with variance $\leq R^2$

(Honest) Confidence Sets

Given the data $((X_1, Y_1), \ldots, (X_n, Y_n))$ and

$$0 \leq \delta \leq 1,$$

construct

$$C_n \subset \mathbb{R}^d$$

such that

$$\mathbf{Pr}(\theta_* \in C_n) \geq 1 - \delta.$$

(Honest) Confidence Sets

Given the data $((X_1, Y_1), \ldots, (X_n, Y_n))$ and

$$0 \le \delta \le 1,$$

construct

$$C_n \subset \mathbb{R}^d$$

such that

$$\mathbf{Pr}(\theta_* \in C_n) \ge 1 - \delta.$$

(Honest) Confidence Sets

Given the data $((X_1, Y_1), \ldots, (X_n, Y_n))$ and

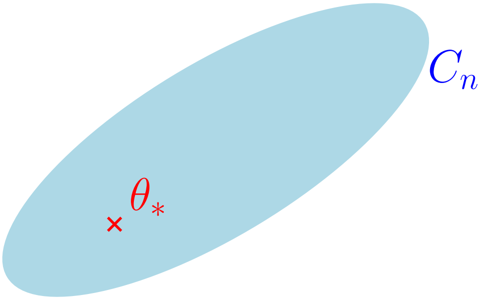
$$0 \leq \delta \leq 1,$$

construct

$$C_n \subset \mathbb{R}^d$$

such that

$$\mathbf{Pr}(\theta_* \in C_n) \geq 1 - \delta.$$



Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \dots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$
- Ridge regression estimator:

$$\widehat{\theta}_n = (\mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I)^{-1}\mathbf{X}_{1:n}^T \mathbf{Y}_{1:n}$$

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$
- Ridge regression estimator:

$$\widehat{\theta}_n = (\mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I)^{-1}\mathbf{X}_{1:n}^T\mathbf{Y}_{1:n}$$

- Let $V_n = \mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I$

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$
- Ridge regression estimator:

$$\widehat{\theta}_n = (\mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I)^{-1}\mathbf{X}_{1:n}^T\mathbf{Y}_{1:n}$$

- Let $V_n = \mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I$

Theorem ([AYPS11])

If $\|\theta_*\|_2 \leq S$, then with probability at least $1 - \delta$, for all t, θ_* lies in

$$C_t = \left\{ \theta \; : \; \|\widehat{\theta}_t - \theta\|_{V_t} \leq R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta \det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda} \right\}$$

where $\|v\|_A = \sqrt{v^T A v}$ is the matrix A-norm.

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$
- Ridge regression estimator:

$$\widehat{\theta}_n = (\mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I)^{-1}\mathbf{X}_{1:n}^T \mathbf{Y}_{1:n}$$

- Let $V_n = \mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I$

Theorem ([AYPS11])

If $\|\theta_*\|_2 \le S$, then with probability at least $1 - \delta$, for all t, θ_* lies in

$$C_t = \left\{ \theta \ : \ \|\widehat{\theta}_t - \theta\|_{V_t} \le R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta \det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda} \right\}$$

where $\|v\|_A = \sqrt{v^T A v}$ is the matrix A-norm.

Proof technique: [RS70, dLS09].

Confidence Sets based on Ridge-regression

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_t \approx \langle X_t, \theta_* \rangle$
- Stack them into matrices: $\mathbf{X}_{1:n}$ is $n \times d$ and $\mathbf{Y}_{1:n}$ is $n \times 1$
- Ridge regression estimator:

$$\widehat{\theta}_n = (\mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I)^{-1}\mathbf{X}_{1:n}^T \mathbf{Y}_{1:n}$$

- Let $V_n = \mathbf{X}_{1:n}\mathbf{X}_{1:n}^T + \lambda I$

Theorem ([AYPS11])

If $\|\theta_*\|_2 \leq S$, then with probability at least $1 - \delta$, for all t, θ_* lies in

$$C_t = \left\{ \theta \ : \ \|\widehat{\theta}_t - \theta\|_{V_t} \leq R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta \det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda} \right\}$$

where $\|v\|_A = \sqrt{v^T A v}$ is the matrix A-norm.

Proof technique: [RS70, dLS09]. Extends to separable Hilbert spaces.

Comparison with Previous Confidence Sets

- Bound of [AYPS11]:

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \le R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta\det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda}$$

Comparison with Previous Confidence Sets

- Bound of [AYPS11]:

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \leq R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta\det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda}$$

- [DHK08]: If $\|\theta_*\|_2, \|X_t\|_2 \leq 1$ then for a specific λ

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \leq R\max\left\{\sqrt{128d\ln(t)\ln(t^2/\delta)}, \frac{8}{3}\ln(t^2/\delta)\right\}$$

Comparison with Previous Confidence Sets

- Bound of [AYPS11]:

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \le R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta\det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda}$$

- [DHK08]: If $\|\theta_*\|_2, \|X_t\|_2 \le 1$ then for a specific λ

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \le R\max\left\{\sqrt{128d\ln(t)\ln(t^2/\delta)}, \frac{8}{3}\ln(t^2/\delta)\right\}$$

- [RT10]: If $\|X_t\|_2 \le 1$

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \le 2R\kappa\sqrt{\ln t}\sqrt{d\ln t + \ln(t^2/\delta)} + S\sqrt{\lambda}$$

where $\kappa = 3 + 2\ln((1+\lambda d)/\lambda)$.

Comparison with Previous Confidence Sets

- Bound of [AYPS11]:

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \leq R\sqrt{2\ln\left(\frac{\det(V_t)^{1/2}}{\delta\det(\lambda I)^{1/2}}\right)} + S\sqrt{\lambda}$$

- [DHK08]: If $\|\theta_*\|_2, \|X_t\|_2 \leq 1$ then for a specific λ

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \leq R\max\left\{\sqrt{128d\ln(t)\ln(t^2/\delta)}, \frac{8}{3}\ln(t^2/\delta)\right\}$$

- [RT10]: If $\|X_t\|_2 \leq 1$

$$\|\widehat{\theta}_t - \theta_*\|_{V_t} \leq 2R\kappa\sqrt{\ln t}\sqrt{d\ln t + \ln(t^2/\delta)} + S\sqrt{\lambda}$$

where $\kappa = 3 + 2\ln((1+\lambda d)/\lambda)$.

The bound of [AYPS11] doesn't depend on t.

Questions

- Are there other ways to construct confidence sets?

Questions

- Are there other ways to construct confidence sets?
- Can we get tighter confidence sets when some special conditions are met?

Questions

- Are there other ways to construct confidence sets?
- Can we get tighter confidence sets when some special conditions are met?
- SPARSITY:

Only p coordinates of θ_* are nonzero.

Questions

- Are there other ways to construct confidence sets?
- Can we get tighter confidence sets when some special conditions are met?
- SPARSITY:

Only p coordinates of θ_* are nonzero.

- Can we construct tighter confidence sets based on the knowledge of p?

Questions

- Are there other ways to construct confidence sets?
- Can we get tighter confidence sets when some special conditions are met?
- SPARSITY:

Only p coordinates of θ_* are nonzero.

- Can we construct tighter confidence sets based on the knowledge of p?
- Least-squares (or ridge) estimators are not a good idea!

Online-to-Confidence-Set Conversion

Online-to-Confidence-Set Conversion

- Idea: Create a confidence set based on how well an online linear prediction algorithm works.

Online-to-Confidence-Set Conversion

- Idea: Create a confidence set based on how well an online linear prediction algorithm works.
- This is a reduction!

Online-to-Confidence-Set Conversion

- Idea: Create a confidence set based on how well an online linear prediction algorithm works.
- This is a reduction!
- If a new prediction algorithm is discovered, or a better performance bounds for an algorithm becomes available, we get tighter confidence sets

Online-to-Confidence-Set Conversion

- Idea: Create a confidence set based on how well an online linear prediction algorithm works.
- This is a reduction!
- If a new prediction algorithm is discovered, or a better performance bounds for an algorithm becomes available, we get tighter confidence sets
- Hopefully it will work for the sparse case

Online-to-Confidence-Set Conversion

- Idea: Create a confidence set based on how well an online linear prediction algorithm works.
- This is a reduction!
- If a new prediction algorithm is discovered, or a better performance bounds for an algorithm becomes available, we get tighter confidence sets
- Hopefully it will work for the sparse case

Encouragement: Working on my thesis

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best <u>linear</u> predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

- online gradient descent [Zin03]

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

- online gradient descent [Zin03]
- online least-squares [AW01, Vov01]

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best <u>linear</u> predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

- online gradient descent [Zin03]
- online least-squares [AW01, Vov01]
- exponentiated gradient algorithm [KW97]

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

- online gradient descent [Zin03]
- online least-squares [AW01, Vov01]
- exponentiated gradient algorithm [KW97]
- online LASSO (??)

Online Linear Prediction

For $t = 1, 2, \ldots$:

- Receive $X_t \in \mathbb{R}^d$
- Predict $\widehat{Y}_t \in \mathbb{R}$
- Receive correct label $Y_t \in \mathbb{R}$
- Suffer loss $(Y_t - \widehat{Y}_t)^2$

Goal: Compete with the best linear predictor in hindsight

No assumptions whatsoever on $(X_1, Y_1), (X_2, Y_2), \ldots$!

There are heaps of algorithms for this problem:

- online gradient descent [Zin03]
- online least-squares [AW01, Vov01]
- exponentiated gradient algorithm [KW97]
- online LASSO (??)
- SeqSEW [Ger11, DT07]

Online Linear Prediction, cnt'd

- Regret with respect to a linear predictor $\theta \in \mathbb{R}^d$

$$\rho_n(\theta) = \sum_{t=1}^{n} (Y_t - \widehat{Y}_t)^2 - \sum_{t=1}^{n} (Y_t - \langle X_t, \theta \rangle)^2$$

Online Linear Prediction, cnt'd

- Regret with respect to a linear predictor $\theta \in \mathbb{R}^d$

$$\rho_n(\theta) = \sum_{t=1}^{n} (Y_t - \widehat{Y}_t)^2 - \sum_{t=1}^{n} (Y_t - \langle X_t, \theta \rangle)^2$$

- Prediction algorithms come with "regret bounds" B_n:

$$\forall n \qquad \rho_n(\theta) \leq B_n$$

Online Linear Prediction, cnt'd

- Regret with respect to a linear predictor $\theta \in \mathbb{R}^d$

$$\rho_n(\theta) = \sum_{t=1}^{n}(Y_t - \widehat{Y}_t)^2 - \sum_{t=1}^{n}(Y_t - \langle X_t, \theta \rangle)^2$$

- Prediction algorithms come with "regret bounds" B_n:

$$\forall n \qquad \rho_n(\theta) \leq B_n$$

- B_n depends on n, d, θ and possibly $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$

Online Linear Prediction, cnt'd

- Regret with respect to a linear predictor $\theta \in \mathbb{R}^d$

$$\rho_n(\theta) = \sum_{t=1}^{n} (Y_t - \widehat{Y}_t)^2 - \sum_{t=1}^{n} (Y_t - \langle X_t, \theta \rangle)^2$$

- Prediction algorithms come with "regret bounds" B_n:

$$\forall n \qquad \rho_n(\theta) \leq B_n$$

- B_n depends on n, d, θ and possibly $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$
- Typically, $B_n = O(\sqrt{n})$ or $B_n = O(\log n)$

Good Regret Implies Small Risk

- Data: $\{(X_t, Y_t)\}_{t=1}^n$ is i.i.d.,
  $Y_t = \langle X_t, \theta_* \rangle + \eta_t$, η_t R-sub-Gaussian

Good Regret Implies Small Risk

- Data: $\{(X_t, Y_t)\}_{t=1}^n$ is i.i.d.,
  $Y_t = \langle X_t, \theta_* \rangle + \eta_t$, η_t R-sub-Gaussian
- Online Learning Algorithm:
  $\mathcal{A}$ produces $\{\theta_t\}_{t=1}^n$ and predicts $\hat{Y}_t = \langle X_t, \theta_t \rangle$

Good Regret Implies Small Risk

- Data: $\{(X_t, Y_t)\}_{t=1}^n$ is i.i.d.,
  $Y_t = \langle X_t, \theta_* \rangle + \eta_t$, η_t R-sub-Gaussian
- Online Learning Algorithm:
  $\mathcal{A}$ produces $\{\theta_t\}_{t=1}^n$ and predicts $\hat{Y}_t = \langle X_t, \theta_t \rangle$
- Regret bound: $\forall n$: $\rho_n(\theta_*) \le B_n$

Good Regret Implies Small Risk

- Data: $\{(X_t, Y_t)\}_{t=1}^n$ is i.i.d.,
  $Y_t = \langle X_t, \theta_* \rangle + \eta_t$, η_t R-sub-Gaussian
- Online Learning Algorithm:
  $\mathcal{A}$ produces $\{\theta_t\}_{t=1}^n$ and predicts $\hat{Y}_t = \langle X_t, \theta_t \rangle$
- Regret bound: $\forall n$: $\rho_n(\theta_*) \le B_n$
- Risk of vector θ: $R(\theta) = \mathbb{E}[(Y_1 - \langle X_1, \theta \rangle)^2]$.

Good Regret Implies Small Risk

- Data: $\{(X_t, Y_t)\}_{t=1}^n$ is i.i.d.,
  $Y_t = \langle X_t, \theta_* \rangle + \eta_t$, η_t R-sub-Gaussian
- Online Learning Algorithm:
  $\mathcal{A}$ produces $\{\theta_t\}_{t=1}^n$ and predicts $\hat{Y}_t = \langle X_t, \theta_t \rangle$
- Regret bound: $\forall n$: $\rho_n(\theta_*) \le B_n$
- Risk of vector θ: $R(\theta) = \mathbf{E}[(Y_1 - \langle X_1, \theta \rangle)^2]$.

Theorem ([CBG08])

Let $\bar{\theta}_n = \frac{1}{n} \sum_{t=1}^n \theta_t$. *Then, w.p.* $1 - \delta$,

$$R(\bar{\theta}_n) \le \frac{B_n}{n} + \frac{36}{n} \ln\left(\frac{B_n + 3}{\delta}\right) + \sqrt{\frac{B_n}{n^2} \ln\left(\frac{B_n + 3}{\delta}\right)}.$$

Online-to-Confidence-Set Conversion

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ where $Y_t = \langle X_t, \theta_* \rangle + \eta_t$ and η_t is conditionally R-sub-Gaussian.

Online-to-Confidence-Set Conversion

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ where $Y_t = \langle X_t, \theta_* \rangle + \eta_t$ and η_t is conditionally R-sub-Gaussian.
- Predictions $\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n$

Online-to-Confidence-Set Conversion

- Data $(X_1, Y_1), \dots, (X_n, Y_n)$ where $Y_t = \langle X_t, \theta_* \rangle + \eta_t$ and η_t is conditionally R-sub-Gaussian.
- Predictions $\widehat{Y}_1, \widehat{Y}_2, \dots, \widehat{Y}_n$
- Regret bound $\rho(\theta_*) \le B_n$

Online-to-Confidence-Set Conversion

- Data $(X_1, Y_1), \ldots, (X_n, Y_n)$ where $Y_t = \langle X_t, \theta_* \rangle + \eta_t$ and η_t is conditionally R-sub-Gaussian.
- Predictions $\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n$
- Regret bound $\rho(\theta_*) \leq B_n$

Theorem (Conversion, [AYPS12])

With probability at least $1 - \delta$, for *all* *n, θ_* lies in*

$$C_n = \left\{ \theta \in \mathbb{R}^d \ : \ \sum_{t=1}^{n} (\hat{Y}_t - \langle X_t, \theta \rangle)^2 \leq 1 + 2B_n + 32R^2 \ln \left(\frac{R\sqrt{8} + \sqrt{1 + B_n}}{\delta} \right) \right\}$$

Proof Sketch

Algebra: With probability 1, due to the regret bound B_n,

$$\sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2 \leq B_n + 2 \underbrace{\sum_{t=1}^{n} \eta_t (\widehat{Y}_t - \langle X_t, \theta_* \rangle)}_{M_n} . \qquad (1)$$

Proof Sketch

Algebra: With probability 1, due to the regret bound B_n,

$$\sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2 \leq B_n + 2 \underbrace{\sum_{t=1}^{n} \eta_t (\widehat{Y}_t - \langle X_t, \theta_* \rangle)}_{M_n} . \qquad (1)$$

$(M_n)_{n=1}^{\infty}$ is a martingale. Using the same argument as in [AYPS11], we get that w.p. $1-\delta$, for all $n \geq 0$,

$$|M_n| \leq R \sqrt{2 \left(1 + \sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2 \right) \times \ln \left(\frac{\sqrt{1 + \sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2}}{\delta} \right)} .$$

Proof Sketch

Algebra: With probability 1, due to the regret bound B_n,

$$\sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2 \le B_n + 2 \underbrace{\sum_{t=1}^{n} \eta_t (\widehat{Y}_t - \langle X_t, \theta_* \rangle)}_{M_n} . \qquad (1)$$

$(M_n)_{n=1}^{\infty}$ is a martingale. Using the same argument as in [AYPS11], we get that w.p. $1 - \delta$, for all $n \ge 0$,

$$|M_n| \le R \sqrt{2 \left(1 + \sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2 \right) \times \ln \left(\frac{\sqrt{1 + \sum_{t=1}^{n} (\widehat{Y}_t - \langle X_t, \theta_* \rangle)^2}}{\delta} \right)} .$$

Combine with (1) and solve the inequality.

Application to Sparse Linear Prediction

Theorem ([Ger11])

For any θ *such that* $\|\theta\|_\infty \le 1$ *and* $\|\theta\|_0 \le p$*, the the regret of* SEQSEW *is bounded by*

$$\rho_n(\theta) \le B_n = O(p \log(nd)) \ .$$

Application to Sparse Linear Prediction

Theorem ([Ger11])

For any θ *such that* $\|\theta\|_\infty \le 1$ *and* $\|\theta\|_0 \le p$*, the the regret of* SEQSEW *is bounded by*

$$\rho_n(\theta) \le B_n = O(p \log(nd)) \ .$$

Corollary

$\exists A > 0$ *s.t. with probability at least* $1 - \delta$*, for all* n*,* θ_* *lies in*

$$C_n = \left\{ \theta \in \mathbb{R}^d \ : \ \sum_{t=1}^{n} (\hat{Y}_t - \langle X_t, \theta \rangle)^2 \le 1 + 2Ap\log(nd) + 32R^2 \ln\left(\frac{R\sqrt{8} + \sqrt{1Ap\log(nd)}}{\delta} \right) \right\} \ .$$

Application to Linear Bandits

Encouragement: Gittin's mistake
Stamina: András' theory of hiking

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

- Receive convex set $D_t \subset \mathbb{R}^d$

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

- Receive convex set $D_t \subset \mathbb{R}^d$
- Choose an action $X_t \in D_t$

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

- Receive convex set $D_t \subset \mathbb{R}^d$
- Choose an action $X_t \in D_t$
- Receive a reward

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t,$$

where η_t is conditionally on the past R-sub-Gaussian.

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

- Receive convex set $D_t \subset \mathbb{R}^d$
- Choose an action $X_t \in D_t$
- Receive a reward

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t,$$

where η_t is conditionally on the past R-sub-Gaussian.

Goal: Maximize total reward.

Linear Bandits

Unknown, but fixed weight vector $\theta_* \in \mathbb{R}^d$.

In round $t = 1, 2, \ldots$

- Receive convex set $D_t \subset \mathbb{R}^d$
- Choose an action $X_t \in D_t$
- Receive a reward

$$Y_t = \langle X_t, \theta_* \rangle + \eta_t,$$

where η_t is conditionally on the past R-sub-Gaussian.

Goal: Maximize total reward.
Sparse bandits: θ_* is sparse.

Motivation

- multi-armed bandits, with infinitely many arms

Motivation

- multi-armed bandits, with infinitely many arms
- dependency between the rewards of arms (linear structure)

Motivation

- multi-armed bandits, with infinitely many arms
- dependency between the rewards of arms (linear structure)
- applications: web-advertisement, network routing, . . .

Motivation

- multi-armed bandits, with infinitely many arms
- dependency between the rewards of arms (linear structure)
- applications: web-advertisement, network routing, ...
  - action = arm = ad = feature vector

Motivation

- multi-armed bandits, with infinitely many arms
- dependency between the rewards of arms (linear structure)
- applications: web-advertisement, network routing, ...
  - action = arm = ad = feature vector
  - reward = click

Motivation

- multi-armed bandits, with infinitely many arms
- dependency between the rewards of arms (linear structure)
- applications: web-advertisement, network routing, ...
  - action = arm = ad = feature vector
  - reward = click
  - sparsity: high-dimensional parameter spaces/feature vectors

Regret

- If we knew θ_*, then in round t we'd choose action

$$X_t^* = \operatorname*{argmax}_{x \in D_t} \langle x, \theta_* \rangle$$

Regret

- If we knew θ_*, then in round t we'd choose action

$$X_t^* = \operatorname*{argmax}_{x \in D_t} \langle x, \theta_* \rangle$$

- Our regret is how much less total reward we have incurred:

$$\mathrm{Regret}_n = \sum_{t=1}^{n} \langle X_t^*, \theta_* \rangle - \sum_{t=1}^{n} \langle X_t, \theta_* \rangle$$

Regret

- If we knew θ_*, then in round t we'd choose action

$$X_t^* = \operatorname*{argmax}_{x \in D_t} \langle x, \theta_* \rangle$$

- Our regret is how much less total reward we have incurred:

$$\text{Regret}_n = \sum_{t=1}^{n} \langle X_t^*, \theta_* \rangle - \sum_{t=1}^{n} \langle X_t, \theta_* \rangle$$

- We want $\text{Regret}_n / n \to 0$ as $n \to \infty$

Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.

Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.
- OFUL Algorithm: In round t, choose

$$(X_t, \widetilde{\theta}_t) = \underset{(x,\theta)\in D_t \times C_{t-1}}{\operatorname{argmax}} \langle X_t, \theta_t \rangle$$

Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.
- OFUL Algorithm: In round t, choose

$$(X_t, \widetilde{\theta}_t) = \underset{(x,\theta)\in D_t \times C_{t-1}}{\operatorname{argmax}} \langle X_t, \theta_t \rangle$$

- $\widetilde{\theta}_t$ is an "optimistic" estimate of θ_*

Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.
- OFUL Algorithm: In round t, choose

$$(X_t, \widetilde{\theta}_t) = \operatorname*{argmax}_{(x,\theta) \in D_t \times C_{t-1}} \langle X_t, \theta_t \rangle$$

- $\widetilde{\theta}_t$ is an "optimistic" estimate of θ_*
- The "OFU" principle goes back to at least [LR85]

Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.
- OFUL Algorithm: In round t, choose

$$(X_t, \widetilde{\theta}_t) = \underset{(x,\theta) \in D_t \times C_{t-1}}{\operatorname{argmax}} \langle X_t, \theta_t \rangle$$

- $\widetilde{\theta}_t$ is an "optimistic" estimate of θ_*
- The "OFU" principle goes back to at least [LR85]
- Algorithm UCB1 of [Aue03] is a special case

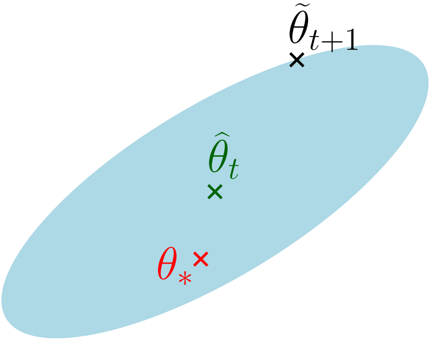
Optimism in the Face of Uncertainty

- Maintain a confidence set $C_t \subseteq \mathbb{R}^d$ such that $\theta_* \in C_t$ with high probability.
- OFUL Algorithm: In round t, choose

$$(X_t, \widetilde{\theta}_t) = \underset{(x,\theta)\in D_t \times C_{t-1}}{\operatorname{argmax}} \langle X_t, \theta_t \rangle$$

- $\widetilde{\theta}_t$ is an "optimistic" estimate of θ_*
- The "OFU" principle goes back to at least [LR85]
- Algorithm UCB1 of [Aue03] is a special case
- Widely applied, very active in machine learning (internet giants)

Confidence Set C_t



- $\widehat{\theta}_t$: center of C_t (e.g., least-squares estimate)
- θ_* lies somewhere in C_t w.h.p.
- Next optimistic estimate, $\widetilde{\theta}_{t+1}$, is on the boundary of C_t

Regret of OFUL with Ridge-Regressor Estimator

Theorem ([DHK08, AYPS11])

If $\|\theta_*\|_2 \leq 1$ *and* D_t*'s are subsets of the unit* 2*-ball then with probability at least* $1 - \delta$

$$\mathrm{Regret}_n \leq O(Rd\sqrt{n} \cdot \mathrm{polylog}(n, d, 1/\delta))$$

Empirical Results

OFUL using the confidence set of [AYPS11]

Empirical Results

OFUL using the confidence set of [AYPS11]
OFUL using the confidence set of [DHK08]

Empirical Results

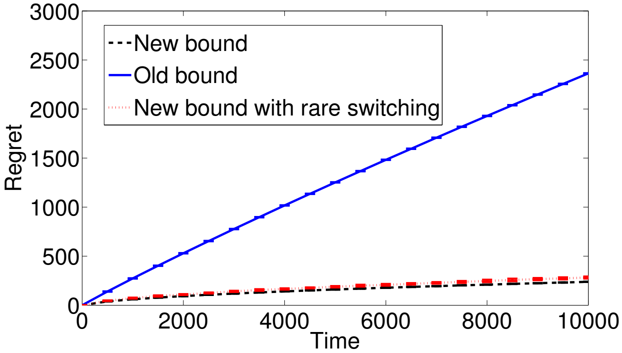
OFUL using the confidence set of [AYPS11] – "New bound"
OFUL using the confidence set of [DHK08]

Empirical Results

OFUL using the confidence set of [AYPS11] – "New bound"
OFUL using the confidence set of [DHK08] – "Old bound"

Empirical Results

OFUL using the confidence set of [AYPS11] – "New bound"
OFUL using the confidence set of [DHK08] – "Old bound"



OFUL with Online-to-Confidence-Set Conversion

Theorem

If $|\langle x, \theta_* \rangle| \leq 1$ *for all* $x \in D_t$ *and all* t *then with probability at least* $1 - \delta$*, for all* n*, the regret of Optimistic Algorithm is*

$$\text{Regret}_n \leq O\left(\sqrt{dnB_n} \cdot \text{polylog}(n, d, 1/\delta, B_n)\right) .$$

OFUL Combined with SeqSEW

Theorem ([AYPS12])

Suppose $\|\theta_*\|_2 \leq 1$ *and* $\|\theta_*\|_0 \leq p$*. Via the conversion, OFUL has regret*

$$O(R\sqrt{pdn} \cdot polylog(n, d, 1/\delta))$$

OFUL Combined with SeqSEW

Theorem ([AYPS12])

Suppose $\|\theta_*\|_2 \leq 1$ *and* $\|\theta_*\|_0 \leq p$. *Via the conversion, OFUL has regret*

$$O(R\sqrt{pdn} \cdot polylog(n, d, 1/\delta))$$

... which is better than $O(Rd\sqrt{n} \cdot \mathrm{polylog}(n, d, 1/\delta))$.

Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm

Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm
OFUL-LS: OFUL with ridge regression

Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm

OFUL-LS: OFUL with ridge regression

$d = 200$: $D_t = D = \mathrm{co}(A_1, \ldots, A_{200})$, $A_i \in \{-1, +1\}^{200}$

Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm

OFUL-LS: OFUL with ridge regression

$d = 200$: $D_t = D = \mathrm{co}(A_1, \ldots, A_{200})$, $A_i \in \{-1, +1\}^{200}$

$p = 10$ with $\theta_{*,i} \in \{0, 0.1\}$,

Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm

OFUL-LS: OFUL with ridge regression

$d = 200$: $D_t = D = \mathrm{co}(A_1, \ldots, A_{200})$, $A_i \in \{-1, +1\}^{200}$

$p = 10$ with $\theta_{*,i} \in \{0, 0.1\}$, $\eta_t \sim \mathcal{N}(0, 0.1^2)$

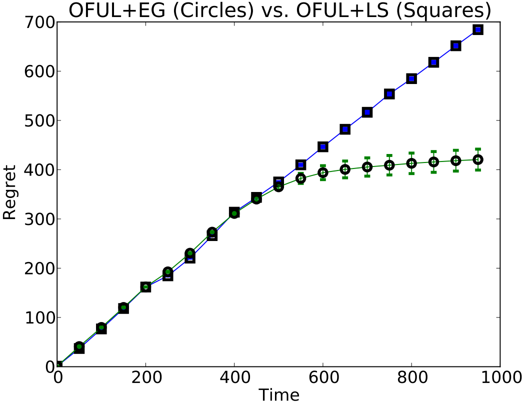
Results on Sparse Bandits

OFUL-EG: OFUL with the EG algorithm

OFUL-LS: OFUL with ridge regression

$d = 200$: $D_t = D = \mathrm{co}(A_1, \ldots, A_{200})$, $A_i \in \{-1, +1\}^{200}$

$p = 10$ with $\theta_{*,i} \in \{0, 0.1\}$, $\eta_t \sim \mathcal{N}(0, 0.1^2)$



Summary

Summary

- Online-to-Confidence-Set Conversion

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits
- Top results on the recent Yahoo article recommendation competition (with no tuning, RR)

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits
- Top results on the recent Yahoo article recommendation competition (with no tuning, RR)

Open Problems

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits
- Top results on the recent Yahoo article recommendation competition (with no tuning, RR)

Open Problems

- Confidence sets for batch algorithms e.g. offline LASSO.

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits
- Top results on the recent Yahoo article recommendation competition (with no tuning, RR)

Open Problems

- Confidence sets for batch algorithms e.g. offline LASSO.
- Adaptive bandit algorithm that doesn't need p upfront.

Summary

- Online-to-Confidence-Set Conversion
- First confidence set for sparse linear prediction
- Application to bandits
- Top results on the recent Yahoo article recommendation competition (with no tuning, RR)

Open Problems

- Confidence sets for batch algorithms e.g. offline LASSO.
- Adaptive bandit algorithm that doesn't need p upfront.
- When D_t has few corners, LS wins. Best of both worlds?

Lesson about health

Lesson about health
Wonderful memories

Lesson about health
Wonderful memories
THANK YOU!

References

[Aue03] P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3:397–422, 2003.

[AW01] Katy S. Azoury and Mafred K. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43:211–246, 2001.

[AYPS11] Y. Abbasi-Yadkori, D. Pál, and Cs. Szepesvári. Improved algorithms for linear stochastic bandits (extended version). In *NIPS-24*, pages 2312–2320. Curran Associates, December 2011.

[AYPS12] Y. Abbasi-Yadkori, D. Pál, and Cs. Szepesvári. Online-to-confidence-set conversions and application to sparse stochastic bandits. In N. Lawrence and M. Girolami, editors, *AISTATS 2012*, volume 22, pages 1–9, 2012.

[CBG08] N. Cesa-Bianchi and C. Gentile. Improved risk tail bounds for on-line algorithms. *IEEE Transactions on Information Theory*, 54(1):386–390, 2008.

[DHK08] Varsha Dani, Thomas P. Hayes, and Sham M. Kakade. Stochastic linear optimization under bandit feedback. In Rocco Servedio and Tong Zhang, editors, *Proceedings of the 21st Annual Conference on Learning Theory (COLT 2008)*, pages 355–366, 2008.

[DT07] Arnak S. Dalalyan and Alexandre B. Tsybakov. Aggregation by exponential weighting and sharp oracle inequalities. In *Proceedings of the 20th Annual Conference on Learning Theory*, pages 97–111, 2007.

[Ger11] Sèbastien Gerchinovitz. Sparsity regret bounds for individual sequences in online linear regression. In *Proceedings of the 24st Annual Conference on Learning Theory (COLT 2011)*, 2011.

[KW97] Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1–63, January 1997.

[LR85] T. L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6:4–22, 1985.

[RS70] H. Robbins and D. Siegmund. Boundary crossing probabilities for the Wiener process and sample sums. *Annals of Math. Statistics*, 41:1410–1429, 1970.

[RT10] Paat Rusmevichientong and John N. Tsitsiklis. Linearly parameterized bandits. *Mathematics of Operations Research*, 35(2):395–411, 2010.

[Vov01] Vladimir Vovk. Competitive on-line statistics. *International Statistical Review*, 69:213–248, 2001.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal