CSE 541: Interactive Learning

Lecture 10: Linear bandits II: elimination and LinUCB

Kevin Jamieson

Paul G. Allen School of
Computer Science & Engineering
University of Washington

Monday, November 2, 2026

Last time, 1: least squares and its error in a direction

Arms $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$, unknown θ^*, pull x_t and see $y_t = \langle x_t, \theta^* \rangle + \eta_t$ with $\eta_t \sim \mathcal{N}(0, 1)$. For a design $x_1, \ldots, x_\tau$ chosen *in advance* (stacked as X, $Y = X\theta^* + \eta$):

$$\hat{\theta} = \arg\min_\theta \|Y - X\theta\|_2^2 = (X^\top X)^{-1} X^\top Y = \theta^* + (X^\top X)^{-1} X^\top \eta.$$

- In a direction z: $\langle z, \hat{\theta} - \theta^* \rangle = \langle w, \eta \rangle$ with $w = X(X^\top X)^{-1} z$ fixed, so it is $\mathcal{N}\big(0, \|z\|^2_{(X^\top X)^{-1}}\big)$ and

$$\mathbb{P}\Big(\langle z, \hat{\theta} - \theta^* \rangle > \|z\|_{(X^\top X)^{-1}} \sqrt{2 \log(1/\delta)}\Big) \le \delta.$$

- Write $V = X^\top X = \sum_t x_t x_t^\top$. With $V = USU^\top$, $\|z\|^2_{V^{-1}} = \sum_j \langle z, u_j \rangle^2 / s_j$: the directions you sampled are the directions you know, and $\{\theta : \left\|\theta - \hat{\theta}\right\|_V \le \beta\}$ is an ellipsoid, tight along well-sampled axes.
- The design must not depend on the noise. The Gaussian (or Chernoff) computation treats w as a constant; an algorithm that picks x_t based on $y_1, \ldots, y_{t-1}$ breaks it.

Last time, 2: experimental design

A design $\lambda \in \triangle_{\mathcal{X}}$ has $A(\lambda) = \sum_x \lambda_x x x^\top$; pulling x about $\lambda_x \tau$ times gives $X^\top X \approx \tau A(\lambda)$, so $\|z\|^2_{(X^\top X)^{-1}} \approx \|z\|^2_{A(\lambda)^{-1}} / \tau$. G criterion: $f(\lambda) = \max_x \|x\|^2_{A(\lambda)^{-1}}$, the squared width of the worst arm.

Kiefer–Wolfowitz

$\min_\lambda f(\lambda) = d$ for every $\mathcal{X}$ spanning $\mathbb{R}^d$; minimizers of f = maximizers of $\log\det A(\lambda)$; a minimizer with support $\leq d(d+1)/2$ exists.

Proposition

Pull $\lceil \tau \lambda^\star_x \rceil$ times each and run least squares. W.p. $\geq 1 - \delta$, for *all* $x \in \mathcal{X}$ at once: $|\langle x, \hat{\theta} - \theta^* \rangle| \leq \sqrt{2d \log(2n/\delta)/\tau}$.

Versus $\sqrt{2n \log(2n/\delta)/\tau}$ for pulling n arms equally: n became d. Compute $\lambda^\star$ by exponential weights or Frank–Wolfe on $\log\det A(\lambda)$, whose gradient is the leverage.

Last time, 2: experimental design

A design $\lambda \in \triangle_{\mathcal{X}}$ has $A(\lambda) = \sum_x \lambda_x x x^\top$; pulling x about $\lambda_x \tau$ times gives $X^\top X \approx \tau A(\lambda)$, so $\|z\|^2_{(X^\top X)^{-1}} \approx \|z\|^2_{A(\lambda)^{-1}} / \tau$. G criterion: $f(\lambda) = \max_x \|x\|^2_{A(\lambda)^{-1}}$, the squared width of the worst arm.

Kiefer–Wolfowitz

$\min_\lambda f(\lambda) = d$ for every $\mathcal{X}$ spanning $\mathbb{R}^d$; minimizers of f = maximizers of $\log\det A(\lambda)$; a minimizer with support $\leq d(d+1)/2$ exists.

Proposition

Pull $\lceil \tau \lambda^\star_x \rceil$ times each and run least squares. W.p. $\geq 1 - \delta$, for *all* $x \in \mathcal{X}$ at once: $|\langle x, \hat{\theta} - \theta^* \rangle| \leq \sqrt{2d \log(2n/\delta)/\tau}$.

Versus $\sqrt{2n \log(2n/\delta)/\tau}$ for pulling n arms equally: n became d. Compute $\lambda^\star$ by exponential weights or Frank–Wolfe on $\log\det A(\lambda)$, whose gradient is the leverage.

Today. Finish elimination, then LinUCB, which uses *all* its data and needs a confidence *ellipsoid* for adaptively collected data: motivated from the fixed design, used as given, proved next lecture.

Successive elimination with features

The same algorithm, with a design in place of “pull everyone equally”

Elimination for linear bandits

Elimination with G-optimal designs (notes §11.4; Lattimore–Szepesvári Ch. 22)

Input: arms $\mathcal{X}$, $|\mathcal{X}| = n$, confidence δ. **Initialize** $\mathcal{X}_1 = \mathcal{X}$, $\ell = 1$.

while $|\mathcal{X}_\ell| > 1$:

$\epsilon_\ell = 2^{-\ell}$, $\tau_\ell = 2d\,\epsilon_\ell^{-2}\log(4\ell^2 n/\delta)$; G-optimal design $\hat{\lambda}_\ell$ *for the survivors* $\mathcal{X}_\ell$ (support $\le d(d+1)/2$)

pull each $x \in \mathcal{X}_\ell$ exactly $\lceil \hat{\lambda}_{\ell,x}\tau_\ell \rceil$ times; least squares on *this round's* data gives $\hat{\theta}_\ell$

$\mathcal{X}_{\ell+1} = \mathcal{X}_\ell \setminus \{x : \max_{x' \in \mathcal{X}_\ell} \langle x' - x, \hat{\theta}_\ell \rangle > 2\epsilon_\ell\}$; $\ell \leftarrow \ell + 1$

Output $\mathcal{X}_\ell$ (for regret: keep playing the survivor).

Elimination for linear bandits

Elimination with G-optimal designs (notes §11.4; Lattimore–Szepesvári Ch. 22)

Input: arms $\mathcal{X}$, $|\mathcal{X}| = n$, confidence δ. **Initialize** $\mathcal{X}_1 = \mathcal{X}$, $\ell = 1$.
while $|\mathcal{X}_\ell| > 1$:
$\epsilon_\ell = 2^{-\ell}$, $\tau_\ell = 2d\,\epsilon_\ell^{-2}\log(4\ell^2 n/\delta)$; G-optimal design $\hat{\lambda}_\ell$ *for the survivors* $\mathcal{X}_\ell$ (support $\le d(d+1)/2$)
pull each $x \in \mathcal{X}_\ell$ exactly $\lceil \hat{\lambda}_{\ell,x}\tau_\ell \rceil$ times; least squares on *this round's* data gives $\hat{\theta}_\ell$
$\mathcal{X}_{\ell+1} = \mathcal{X}_\ell \setminus \{x : \max_{x' \in \mathcal{X}_\ell} \langle x' - x, \hat{\theta}_\ell \rangle > 2\epsilon_\ell\}$; $\ell \leftarrow \ell + 1$
Output $\mathcal{X}_\ell$ (for regret: keep playing the survivor).

What changed

- $2\epsilon_\ell^{-2}\log(\cdot)$ pulls *per arm* became $2d\epsilon_\ell^{-2}\log(\cdot)$ *in total*, spread by the design.
- "empirical mean of x" became $\langle x, \hat{\theta}_\ell \rangle$, defined even for arms never pulled.

What did not

- precision doubles, samples quadruple, ℓ^2 pays for the union bound over rounds;
- fresh samples each round, so $\hat{\theta}_\ell$ is a *fixed-design* estimate given $\mathcal{X}_\ell$.

The survivor sets are random: handling that honestly

Lemma (notes Lemma 20)

Assume $\max_x \Delta_x \le 4$. W.p. $\ge 1-\delta$, for every ℓ: $x^\star \in \mathcal{X}_\ell$ and $\max_{x \in \mathcal{X}_\ell} \Delta_x \le 8\epsilon_\ell$.

The subtlety. Round ℓ's design, pulls and $\hat{\theta}_\ell$ depend on $\mathcal{X}_\ell$, which is *random*: carved out by the noise of rounds $< \ell$. The proposition was for a fixed arm set, so fix the arm set first.

The survivor sets are random: handling that honestly

Lemma (notes Lemma 20)

Assume $\max_x \Delta_x \le 4$. W.p. $\ge 1-\delta$, for every ℓ: $x^\star \in \mathcal{X}_\ell$ and $\max_{x\in\mathcal{X}_\ell} \Delta_x \le 8\epsilon_\ell$.

The subtlety. Round ℓ's design, pulls and $\hat{\theta}_\ell$ depend on $\mathcal{X}_\ell$, which is *random*: carved out by the noise of rounds $< \ell$. The proposition was for a fixed arm set, so fix the arm set first.

For *fixed* $V \subseteq \mathcal{X}$, let $\hat{\theta}_\ell(V)$ be the round-ℓ estimate if $\mathcal{X}_\ell = V$ (design for V, τ_ℓ fresh samples, least squares), and $E_{x,\ell}(V) = \{|\langle x, \hat{\theta}_\ell(V) - \theta^*\rangle| \le \epsilon_\ell\}$. For fixed V the proposition applies, with $\sqrt{2d\log(4\ell^2 n/\delta)/\tau_\ell} =$ so $\mathbb{P}\big(\bigcup_{x\in V} E^c_{x,\ell}(V)\big) \le \delta/(2\ell^2)$.

The survivor sets are random: handling that honestly

Lemma (notes Lemma 20)

Assume $\max_x \Delta_x \le 4$. W.p. $\ge 1-\delta$, for every ℓ: $x^\star \in \mathcal{X}_\ell$ and $\max_{x \in \mathcal{X}_\ell} \Delta_x \le 8\epsilon_\ell$.

The subtlety. Round ℓ's design, pulls and $\hat{\theta}_\ell$ depend on $\mathcal{X}_\ell$, which is *random*: carved out by the noise of rounds $< \ell$. The proposition was for a fixed arm set, so fix the arm set first.

For *fixed* $V \subseteq \mathcal{X}$, let $\hat{\theta}_\ell(V)$ be the round-ℓ estimate if $\mathcal{X}_\ell = V$ (design for V, τ_ℓ fresh samples, least squares), and $E_{x,\ell}(V) = \{|\langle x, \hat{\theta}_\ell(V) - \theta^* \rangle| \le \epsilon_\ell\}$. For fixed V the proposition applies, with $\sqrt{2d\log(4\ell^2 n/\delta)/\tau_\ell} = \epsilon_\ell$, so $\mathbb{P}\big(\bigcup_{x \in V} E^c_{x,\ell}(V)\big) \le \delta/(2\ell^2)$.

The survivor sets are random: handling that honestly

Lemma (notes Lemma 20)

Assume $\max_x \Delta_x \le 4$. W.p. $\ge 1-\delta$, for every ℓ: $x^\star \in \mathcal{X}_\ell$ and $\max_{x\in\mathcal{X}_\ell} \Delta_x \le 8\epsilon_\ell$.

The subtlety. Round ℓ's design, pulls and $\hat{\theta}_\ell$ depend on $\mathcal{X}_\ell$, which is *random*: carved out by the noise of rounds $< \ell$. The proposition was for a fixed arm set, so fix the arm set first.

For *fixed* $V \subseteq \mathcal{X}$, let $\hat{\theta}_\ell(V)$ be the round-ℓ estimate if $\mathcal{X}_\ell = V$ (design for V, τ_ℓ fresh samples, least squares), and $E_{x,\ell}(V) = \{|\langle x, \hat{\theta}_\ell(V) - \theta^*\rangle| \le \epsilon_\ell\}$. For fixed V the proposition applies, with $\sqrt{2d\log(4\ell^2 n/\delta)/\tau_\ell} = \epsilon_\ell$, so $\mathbb{P}\big(\bigcup_{x\in V} E^c_{x,\ell}(V)\big) \le \delta/(2\ell^2)$.

This is a statement about a fixed set V and the fresh noise of round ℓ. It says nothing yet about the set the algorithm actually has, $\mathcal{X}_\ell$.

Now let the set be random

Two sources of randomness, and they are independent: $\{\mathcal{X}_\ell = V\}$ is determined by the noise of rounds $< \ell$; $E_{x,\ell}(V)$ by the fresh noise of round ℓ. Decompose over the (finitely many) values of $\mathcal{X}_\ell$:

$$\mathbb{P}\Big(\bigcup_{x\in\mathcal{X}_\ell} E^c_{x,\ell}(\mathcal{X}_\ell)\Big) = \sum_{V\subseteq\mathcal{X}} \mathbb{P}\Big(\bigcup_{x\in V} E^c_{x,\ell}(V),\ \mathcal{X}_\ell = V\Big)$$
$$=$$
$$\leq$$

Now let the set be random

Two sources of randomness, and they are independent: $\{\mathcal{X}_\ell = V\}$ is determined by the noise of rounds $< \ell$; $E_{x,\ell}(V)$ by the fresh noise of round ℓ. Decompose over the (finitely many) values of $\mathcal{X}_\ell$:

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{x\in\mathcal{X}_\ell} E^c_{x,\ell}(\mathcal{X}_\ell)\Big) &= \sum_{V\subseteq\mathcal{X}} \mathbb{P}\Big(\bigcup_{x\in V} E^c_{x,\ell}(V),\ \mathcal{X}_\ell = V\Big)\\ &= \sum_{V} \mathbb{P}\Big(\bigcup_{x\in V} E^c_{x,\ell}(V)\Big)\mathbb{P}(\mathcal{X}_\ell = V)\\ &\le \frac{\delta}{2\ell^2}\sum_{V}\mathbb{P}(\mathcal{X}_\ell = V) = \frac{\delta}{2\ell^2}.\end{aligned}$$

Sum over ℓ: the good event $E = \bigcap_\ell \bigcap_{x\in\mathcal{X}_\ell} E_{x,\ell}(\mathcal{X}_\ell)$ has $\mathbb{P}(E) \ge 1-\delta$.

Now let the set be random

Two sources of randomness, and they are independent: $\{\mathcal{X}_\ell = V\}$ is determined by the noise of rounds $< \ell$; $E_{x,\ell}(V)$ by the fresh noise of round ℓ. Decompose over the (finitely many) values of $\mathcal{X}_\ell$:

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{x\in\mathcal{X}_\ell} E^c_{x,\ell}(\mathcal{X}_\ell)\Big) &= \sum_{V\subseteq\mathcal{X}} \mathbb{P}\Big(\bigcup_{x\in V} E^c_{x,\ell}(V),\ \mathcal{X}_\ell = V\Big)\\ &= \sum_V \mathbb{P}\Big(\bigcup_{x\in V} E^c_{x,\ell}(V)\Big)\mathbb{P}(\mathcal{X}_\ell = V)\\ &\le \frac{\delta}{2\ell^2}\sum_V \mathbb{P}(\mathcal{X}_\ell = V) = \frac{\delta}{2\ell^2}.\end{aligned}$$

Sum over ℓ: the good event $E = \bigcap_\ell \bigcap_{x\in\mathcal{X}_\ell} E_{x,\ell}(\mathcal{X}_\ell)$ has $\mathbb{P}(E) \ge 1-\delta$.

Why fresh samples

The step $\mathbb{P}(A, \mathcal{X}_\ell = V) = \mathbb{P}(A)\mathbb{P}(\mathcal{X}_\ell = V)$ is where independence was used. If round ℓ reused earlier data, $\hat{\theta}_\ell$ and $\mathcal{X}_\ell$ would share noise, the factorization would fail, and the fixed-design interval would not apply. That is the price of today's simplicity.

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$.

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$. $x^\star$ **survives.** If $x^\star \in \mathcal{X}_\ell$ then for every $x \in \mathcal{X}_\ell$,

$$\langle x - x^\star, \hat{\theta}_\ell \rangle =$$

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$. $x^\star$ **survives.** If $x^\star \in \mathcal{X}_\ell$ then for every $x \in \mathcal{X}_\ell$,

$$\langle x - x^\star, \hat{\theta}_\ell \rangle = \langle x, \hat{\theta}_\ell - \theta^* \rangle - \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \Delta_x \;\le\; \epsilon_\ell + \epsilon_\ell - \Delta_x \;\le\; 2\epsilon_\ell,$$

so $\max_{x'} \langle x' - x^\star, \hat{\theta}_\ell \rangle \le 2\epsilon_\ell$ and $x^\star \in \mathcal{X}_{\ell+1}$.

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$. $x^\star$ **survives.** If $x^\star \in \mathcal{X}_\ell$ then for every $x \in \mathcal{X}_\ell$,

$$\langle x - x^\star, \hat{\theta}_\ell \rangle = \langle x, \hat{\theta}_\ell - \theta^* \rangle - \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \Delta_x \;\le\; \epsilon_\ell + \epsilon_\ell - \Delta_x \;\le\; 2\epsilon_\ell,$$

so $\max_{x'} \langle x' - x^\star, \hat{\theta}_\ell \rangle \le 2\epsilon_\ell$ and $x^\star \in \mathcal{X}_{\ell+1}$.

Bad arms go. If $x \in \mathcal{X}_\ell$ has $\Delta_x > 4\epsilon_\ell$, then (using $x^\star \in \mathcal{X}_\ell$)

$$\langle x^\star - x, \hat{\theta}_\ell \rangle =$$

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$. $x^\star$ **survives.** If $x^\star \in \mathcal{X}_\ell$ then for every $x \in \mathcal{X}_\ell$,

$$\langle x - x^\star, \hat{\theta}_\ell \rangle = \langle x, \hat{\theta}_\ell - \theta^* \rangle - \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \Delta_x \;\le\; \epsilon_\ell + \epsilon_\ell - \Delta_x \;\le\; 2\epsilon_\ell,$$

so $\max_{x'} \langle x' - x^\star, \hat{\theta}_\ell \rangle \le 2\epsilon_\ell$ and $x^\star \in \mathcal{X}_{\ell+1}$.

Bad arms go. If $x \in \mathcal{X}_\ell$ has $\Delta_x > 4\epsilon_\ell$, then (using $x^\star \in \mathcal{X}_\ell$)

$$\langle x^\star - x, \hat{\theta}_\ell \rangle = \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \langle x, \hat{\theta}_\ell - \theta^* \rangle + \Delta_x \;>\; -2\epsilon_\ell + 4\epsilon_\ell = 2\epsilon_\ell,$$

so $\max_{x' \in \mathcal{X}_\ell} \langle x' - x, \hat{\theta}_\ell \rangle > 2\epsilon_\ell$ and $x \notin \mathcal{X}_{\ell+1}$. Hence every survivor of round ℓ has $\Delta_x \le 4\epsilon_\ell = 8\epsilon_{\ell+1}$ (and $\le 4 = 8\epsilon_1$ at $\ell = 1$). ■

On the good event: the multi-armed proof, verbatim

Assume E: for every ℓ and every surviving x, $|\langle x, \hat{\theta}_\ell - \theta^* \rangle| \le \epsilon_\ell$. Induct on ℓ; $x^\star \in \mathcal{X}_1$. $x^\star$ **survives.** If $x^\star \in \mathcal{X}_\ell$ then for every $x \in \mathcal{X}_\ell$,

$$\langle x - x^\star, \hat{\theta}_\ell \rangle = \langle x, \hat{\theta}_\ell - \theta^* \rangle - \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \Delta_x \;\le\; \epsilon_\ell + \epsilon_\ell - \Delta_x \;\le\; 2\epsilon_\ell,$$

so $\max_{x'} \langle x' - x^\star, \hat{\theta}_\ell \rangle \le 2\epsilon_\ell$ and $x^\star \in \mathcal{X}_{\ell+1}$.

Bad arms go. If $x \in \mathcal{X}_\ell$ has $\Delta_x > 4\epsilon_\ell$, then (using $x^\star \in \mathcal{X}_\ell$)

$$\langle x^\star - x, \hat{\theta}_\ell \rangle = \langle x^\star, \hat{\theta}_\ell - \theta^* \rangle - \langle x, \hat{\theta}_\ell - \theta^* \rangle + \Delta_x \;>\; -2\epsilon_\ell + 4\epsilon_\ell = 2\epsilon_\ell,$$

so $\max_{x' \in \mathcal{X}_\ell} \langle x' - x, \hat{\theta}_\ell \rangle > 2\epsilon_\ell$ and $x \notin \mathcal{X}_{\ell+1}$. Hence every survivor of round ℓ has $\Delta_x \le 4\epsilon_\ell = 8\epsilon_{\ell+1}$ (and $\le 4 = 8\epsilon_1$ at $\ell = 1$). ■

What the geometry of $\mathcal{X}$ did

Nothing, in this part. It entered only through the width ϵ_ℓ, and Kiefer–Wolfowitz made that width the same for every arm set. The two lines above are the multi-armed elimination lemma with "empirical mean" replaced by $\langle x, \hat{\theta}_\ell \rangle$.

Regret: d replaces n

Theorem (notes §11.4)

Assume $\max_x \Delta_x \le 4$, $\Delta = \min_{x \neq x^\star} \Delta_x$. W.p. $\ge 1 - \delta$, after T pulls,

$$\sum_{x \neq x^\star} \Delta_x T_x \lesssim \frac{d}{\Delta} \log \frac{n \log(1/\Delta)}{\delta} \qquad \text{and} \qquad \sum_{x \neq x^\star} \Delta_x T_x \lesssim \sqrt{dT \log \frac{n \log T}{\delta}}.$$

Regret: d replaces n

Theorem (notes §11.4)

Assume $\max_x \Delta_x \le 4$, $\Delta = \min_{x \neq x^\star} \Delta_x$. W.p. $\ge 1 - \delta$, after T pulls,

$$\sum_{x \neq x^\star} \Delta_x T_x \;\lesssim\; \frac{d}{\Delta} \log \frac{n \log(1/\Delta)}{\delta} \qquad \text{and} \qquad \sum_{x \neq x^\star} \Delta_x T_x \;\lesssim\; \sqrt{dT \log \frac{n \log T}{\delta}}.$$

Proof idea. On E, each pull in round ℓ costs at most $8\epsilon_\ell$ (lemma); round ℓ makes at most $\tau_\ell + d(d+1)/2$ pulls; after $L = \lceil \log_2(8/\Delta) \rceil$ rounds only $x^\star$ survives:

$$\sum_{x \neq x^\star} \Delta_x T_x \;\le\; \sum_{\ell=1}^{L} 8\epsilon_\ell \left(\tau_\ell + \tfrac{d(d+1)}{2} \right) \;\le$$

$$\lesssim$$

Regret: d replaces n

Theorem (notes §11.4)

Assume $\max_x \Delta_x \le 4$, $\Delta = \min_{x \neq x^\star} \Delta_x$. W.p. $\ge 1-\delta$, after T pulls,

$$\sum_{x \neq x^\star} \Delta_x T_x \;\lesssim\; \frac{d}{\Delta} \log \frac{n \log(1/\Delta)}{\delta} \qquad \text{and} \qquad \sum_{x \neq x^\star} \Delta_x T_x \;\lesssim\; \sqrt{dT \log \frac{n \log T}{\delta}}.$$

Proof idea. On E, each pull in round ℓ costs at most $8\epsilon_\ell$ (lemma); round ℓ makes at most $\tau_\ell + d(d+1)/2$ pulls; after $L = \lceil \log_2(8/\Delta) \rceil$ rounds only $x^\star$ survives:

$$\begin{aligned} \sum_{x \neq x^\star} \Delta_x T_x \;\le\; \sum_{\ell=1}^{L} 8\epsilon_\ell \Big(\tau_\ell + \tfrac{d(d+1)}{2} \Big) \;&\le\; \sum_{\ell=1}^{L} 16 d\, \epsilon_\ell^{-1} \log \frac{4\ell^2 n}{\delta} + 4d^2 L \\ &\lesssim\; \frac{d}{\Delta} \log \frac{n \log(1/\Delta)}{\delta}, \end{aligned}$$

since $\sum_{\ell \le L} 2^\ell \le 32/\Delta$ (last round dominates). Gap-free: arms with $\Delta_x \le \nu$ cost at most νT, the rest $\lesssim \frac{d}{\nu} \log(\cdot)$; take $\nu \approx \sqrt{d \log(\cdot)/T}$.

Linear UCB

Optimism with a confidence ellipsoid

What elimination leaves on the table

- **It throws data away.** Each round fits $\hat{\theta}_\ell$ to that round's samples only, so that the design is fixed given $\mathcal{X}_\ell$. Everything from earlier rounds is wasted.
- **It needs a finite, fixed arm set** to compute designs over, and $\log n$ appears. Many problems have infinitely many arms, or a different arm set every round (a different user's candidate items).
- **It is not anytime**, and it over-samples arms with large gaps because the design is spread over all survivors.

What elimination leaves on the table

- **It throws data away.** Each round fits $\hat{\theta}_\ell$ to that round's samples only, so that the design is fixed given $\mathcal{X}_\ell$. Everything from earlier rounds is wasted.
- **It needs a finite, fixed arm set** to compute designs over, and $\log n$ appears. Many problems have infinitely many arms, or a different arm set every round (a different user's candidate items).
- **It is not anytime**, and it over-samples arms with large gaps because the design is spread over all survivors.

What we would like

Use all the data: $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$ and $\hat{\theta}_t = V_t^{-1} \sum_{s \le t} x_s y_s$ (ridge regression, $\lambda > 0$ keeps V_t invertible). Then play optimistically, as UCB did: the arm that *could* be best given the uncertainty about θ^*.

What elimination leaves on the table

- **It throws data away.** Each round fits $\hat{\theta}_\ell$ to that round's samples only, so that the design is fixed given $\mathcal{X}_\ell$. Everything from earlier rounds is wasted.
- **It needs a finite, fixed arm set** to compute designs over, and $\log n$ appears. Many problems have infinitely many arms, or a different arm set every round (a different user's candidate items).
- **It is not anytime**, and it over-samples arms with large gaps because the design is spread over all survivors.

What we would like

Use all the data: $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$ and $\hat{\theta}_t = V_t^{-1} \sum_{s \le t} x_s y_s$ (ridge regression, $\lambda > 0$ keeps V_t invertible). Then play optimistically, as UCB did: the arm that *could* be best given the uncertainty about θ^*.

Two questions. What is "the uncertainty about θ^*" when the x_s were chosen adaptively? And once we have it, how do we turn it into an index for each arm? For the first we need a confidence *set* for the whole vector θ^*, not one direction at a time. Start with the fixed design, where we can compute everything.

A confidence ellipsoid, fixed design first: one direction is easy

Fixed design $X \in \mathbb{R}^{\tau \times d}$, $V = X^\top X$ invertible, independent 1-sub-Gaussian noise. From the first recap slide, for any *fixed* $z \in \mathbb{R}^d$,

$$\mathbb{P}\left(\langle z, \hat{\theta} - \theta^* \rangle \geq \|z\|_{V^{-1}} \sqrt{2 \log(1/\delta)} \right) \leq \delta. \qquad (*)$$

A confidence ellipsoid, fixed design first: one direction is easy

Fixed design $X \in \mathbb{R}^{\tau \times d}$, $V = X^\top X$ invertible, independent 1-sub-Gaussian noise. From the first recap slide, for any *fixed* $z \in \mathbb{R}^d$,

$$\mathbb{P}\left(\langle z, \hat{\theta} - \theta^* \rangle \geq \|z\|_{V^{-1}} \sqrt{2 \log(1/\delta)} \right) \leq \delta. \qquad (*)$$

We want the whole error at once, $\left\| \hat{\theta} - \theta^* \right\|_V$. Write it as a maximum over directions:

$$\left\| \hat{\theta} - \theta^* \right\|_V = \max_{\|u\|_2 = 1}$$

A confidence ellipsoid, fixed design first: one direction is easy

Fixed design $X \in \mathbb{R}^{\tau \times d}$, $V = X^\top X$ invertible, independent 1-sub-Gaussian noise. From the first recap slide, for any *fixed* $z \in \mathbb{R}^d$,

$$\mathbb{P}\Big(\langle z, \hat{\theta} - \theta^* \rangle \geq \|z\|_{V^{-1}} \sqrt{2 \log(1/\delta)} \Big) \leq \delta. \qquad (*)$$

We want the whole error at once, $\left\| \hat{\theta} - \theta^* \right\|_V$. Write it as a maximum over directions:

$$\left\| \hat{\theta} - \theta^* \right\|_V = \max_{\|u\|_2 = 1} \langle V^{1/2} u, \hat{\theta} - \theta^* \rangle, \qquad \text{attained at} \quad U := \frac{V^{1/2}(\hat{\theta} - \theta^*)}{\left\| \hat{\theta} - \theta^* \right\|_V}.$$

For each fixed u on the sphere, $(*)$ with $z = V^{1/2} u$ and $\left\| V^{1/2} u \right\|_{V^{-1}} = \|u\|_2 = 1$ gives $\langle V^{1/2} u, \hat{\theta} - \theta^* \rangle \leq \sqrt{2 \log(1/\delta)}$ w.p. $1 - \delta$.

A confidence ellipsoid, fixed design first: one direction is easy

Fixed design $X \in \mathbb{R}^{\tau \times d}$, $V = X^\top X$ invertible, independent 1-sub-Gaussian noise. From the first recap slide, for any *fixed* $z \in \mathbb{R}^d$,

$$\mathbb{P}\Big(\langle z, \hat{\theta} - \theta^* \rangle \geq \|z\|_{V^{-1}} \sqrt{2\log(1/\delta)}\Big) \leq \delta. \qquad (*)$$

We want the whole error at once, $\left\|\hat{\theta} - \theta^*\right\|_V$. Write it as a maximum over directions:

$$\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2 = 1} \langle V^{1/2} u, \hat{\theta} - \theta^* \rangle, \qquad \text{attained at} \quad U := \frac{V^{1/2}(\hat{\theta} - \theta^*)}{\left\|\hat{\theta} - \theta^*\right\|_V}.$$

For each fixed u on the sphere, $(*)$ with $z = V^{1/2}u$ and $\left\|V^{1/2}u\right\|_{V^{-1}} = \|u\|_2 = 1$ gives $\langle V^{1/2}u, \hat{\theta} - \theta^* \rangle \leq \sqrt{2\log(1/\delta)}$ w.p. $1 - \delta$.

The problem

The maximizer U is *random*: it is built from the noise. $(*)$ applies to fixed directions only, and there are infinitely many directions to union bound over.

The fix is the standard one: a union bound over a *finite net* of directions, plus a little slack to reach every other direction from the net.

Covering the sphere

Lemma (Lattimore–Szepesvári Lemma 20.1)

For every $\epsilon \in (0, 1)$ there is a set $\mathcal{C}_\epsilon \subset S^{d-1} = \{u : \|u\|_2 = 1\}$ with $|\mathcal{C}_\epsilon| \leq (3/\epsilon)^d$ such that every $u \in S^{d-1}$ has some $y \in \mathcal{C}_\epsilon$ with $\|u - y\|_2 \leq \epsilon$.

Covering the sphere

Lemma (Lattimore–Szepesvári Lemma 20.1)

For every $\epsilon \in (0, 1)$ there is a set $\mathcal{C}_\epsilon \subset S^{d-1} = \{u : \|u\|_2 = 1\}$ with $|\mathcal{C}_\epsilon| \leq (3/\epsilon)^d$ such that every $u \in S^{d-1}$ has some $y \in \mathcal{C}_\epsilon$ with $\|u - y\|_2 \leq \epsilon$.

Why. Take a maximal set of points on the sphere that are pairwise more than ϵ apart (it is a cover: any uncovered point could be added). The balls of radius $\epsilon/2$ around them are disjoint and sit inside the ball of radius $1 + \epsilon/2$, so comparing volumes,
$|\mathcal{C}_\epsilon| \, (\epsilon/2)^d \leq$

Covering the sphere

Lemma (Lattimore–Szepesvári Lemma 20.1)

For every $\epsilon \in (0, 1)$ there is a set $\mathcal{C}_\epsilon \subset S^{d-1} = \{u : \|u\|_2 = 1\}$ with $|\mathcal{C}_\epsilon| \leq (3/\epsilon)^d$ such that every $u \in S^{d-1}$ has some $y \in \mathcal{C}_\epsilon$ with $\|u - y\|_2 \leq \epsilon$.

Why. Take a maximal set of points on the sphere that are pairwise more than ϵ apart (it is a cover: any uncovered point could be added). The balls of radius $\epsilon/2$ around them are disjoint and sit inside the ball of radius $1 + \epsilon/2$, so comparing volumes, $|\mathcal{C}_\epsilon| \, (\epsilon/2)^d \leq (1 + \epsilon/2)^d$, i.e. $|\mathcal{C}_\epsilon| \leq (1 + 2/\epsilon)^d \leq (3/\epsilon)^d$.

Covering the sphere

Lemma (Lattimore–Szepesvári Lemma 20.1)

For every $\epsilon \in (0, 1)$ there is a set $\mathcal{C}_\epsilon \subset S^{d-1} = \{u : \|u\|_2 = 1\}$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ such that every $u \in S^{d-1}$ has some $y \in \mathcal{C}_\epsilon$ with $\|u - y\|_2 \le \epsilon$.

Why. Take a maximal set of points on the sphere that are pairwise more than ϵ apart (it is a cover: any uncovered point could be added). The balls of radius $\epsilon/2$ around them are disjoint and sit inside the ball of radius $1 + \epsilon/2$, so comparing volumes, $|\mathcal{C}_\epsilon|\,(\epsilon/2)^d \le (1 + \epsilon/2)^d$, i.e. $|\mathcal{C}_\epsilon| \le (1 + 2/\epsilon)^d \le (3/\epsilon)^d$.

Union bound over the net. Apply $(*)$ to each of the $|\mathcal{C}_\epsilon|$ fixed directions $z = V^{1/2}y$ at level $\delta/|\mathcal{C}_\epsilon|$. With probability at least $1 - \delta$,

for all $y \in \mathcal{C}_\epsilon$: $\quad \langle V^{1/2}y, \hat{\theta} - \theta^* \rangle \le$

Covering the sphere

Lemma (Lattimore–Szepesvári Lemma 20.1)

For every $\epsilon \in (0,1)$ there is a set $\mathcal{C}_\epsilon \subset S^{d-1} = \{u : \|u\|_2 = 1\}$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ such that every $u \in S^{d-1}$ has some $y \in \mathcal{C}_\epsilon$ with $\|u - y\|_2 \le \epsilon$.

Why. Take a maximal set of points on the sphere that are pairwise more than ϵ apart (it is a cover: any uncovered point could be added). The balls of radius $\epsilon/2$ around them are disjoint and sit inside the ball of radius $1 + \epsilon/2$, so comparing volumes, $|\mathcal{C}_\epsilon|\,(\epsilon/2)^d \le (1+\epsilon/2)^d$, i.e. $|\mathcal{C}_\epsilon| \le (1+2/\epsilon)^d \le (3/\epsilon)^d$.

Union bound over the net. Apply $(*)$ to each of the $|\mathcal{C}_\epsilon|$ fixed directions $z = V^{1/2}y$ at level $\delta/|\mathcal{C}_\epsilon|$. With probability at least $1-\delta$,

$$\text{for all } y \in \mathcal{C}_\epsilon : \qquad \langle V^{1/2}y, \hat{\theta} - \theta^* \rangle \;\le\; \sqrt{2\log\frac{|\mathcal{C}_\epsilon|}{\delta}} \;\le\; \sqrt{2d\log\frac{3}{\epsilon} + 2\log\frac{1}{\delta}}.$$

The net has $e^{O(d)}$ points, so the union bound costs $\sqrt{d}$, not $\sqrt{\log n}$. This is where the $\sqrt{d}$ in the radius comes from.

From the net to every direction

On the good event of the last slide, let U be the (random) maximizing direction and pick $y \in \mathcal{C}_\epsilon$ with $\|U - y\|_2 \leq \epsilon$. Then

$$\left\|\hat{\theta} - \theta^*\right\|_V = \langle V^{1/2}U, \hat{\theta} - \theta^* \rangle = \langle V^{1/2}y, \hat{\theta} - \theta^* \rangle + \langle V^{1/2}(U - y), \hat{\theta} - \theta^* \rangle$$
$$\leq$$
$$\leq$$

From the net to every direction

On the good event of the last slide, let U be the (random) maximizing direction and pick $y \in \mathcal{C}_\epsilon$ with $\|U - y\|_2 \le \epsilon$. Then

$$\begin{aligned}\left\|\hat{\theta} - \theta^*\right\|_V = \langle V^{1/2}U, \hat{\theta} - \theta^* \rangle &= \langle V^{1/2}y, \hat{\theta} - \theta^* \rangle + \langle V^{1/2}(U - y), \hat{\theta} - \theta^* \rangle \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \|U - y\|_2 \left\|\hat{\theta} - \theta^*\right\|_V \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \epsilon \left\|\hat{\theta} - \theta^*\right\|_V .\end{aligned}$$

(Middle step: Cauchy–Schwarz, $\langle V^{1/2}w, v\rangle \le \left\|V^{1/2}w\right\|_{V^{-1}} \|v\|_V = \|w\|_2 \|v\|_V$.) Rearrange, and take $\epsilon = 1/2$:

From the net to every direction

On the good event of the last slide, let U be the (random) maximizing direction and pick $y \in \mathcal{C}_\epsilon$ with $\|U - y\|_2 \le \epsilon$. Then

$$\begin{aligned}\left\|\hat{\theta} - \theta^*\right\|_V = \langle V^{1/2}U, \hat{\theta} - \theta^* \rangle &= \langle V^{1/2}y, \hat{\theta} - \theta^* \rangle + \langle V^{1/2}(U - y), \hat{\theta} - \theta^* \rangle \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \|U - y\|_2 \left\|\hat{\theta} - \theta^*\right\|_V \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \epsilon \left\|\hat{\theta} - \theta^*\right\|_V .\end{aligned}$$

(Middle step: Cauchy–Schwarz, $\langle V^{1/2}w, v\rangle \le \left\|V^{1/2}w\right\|_{V^{-1}} \|v\|_V = \|w\|_2 \|v\|_V$.) Rearrange, and take $\epsilon = 1/2$:

Fixed-design confidence ellipsoid (Lattimore–Szepesvári, eq. 20.3)

With probability at least $1 - \delta$:

$$\left\|\hat{\theta} - \theta^*\right\|_V \;\le\; \frac{1}{1-\epsilon}\sqrt{2\log\frac{|\mathcal{C}_\epsilon|}{\delta}} \;\le\; 2\sqrt{2d\log 6 + 2\log\tfrac{1}{\delta}} \;=:\; \beta_{\text{fixed}}.$$

From the net to every direction

On the good event of the last slide, let U be the (random) maximizing direction and pick $y \in \mathcal{C}_\epsilon$ with $\|U - y\|_2 \le \epsilon$. Then

$$\begin{aligned}\left\|\hat{\theta} - \theta^*\right\|_V = \langle V^{1/2}U, \hat{\theta} - \theta^*\rangle &= \langle V^{1/2}y, \hat{\theta} - \theta^*\rangle + \langle V^{1/2}(U-y), \hat{\theta} - \theta^*\rangle \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \|U - y\|_2 \left\|\hat{\theta} - \theta^*\right\|_V \\ &\le \sqrt{2\log\tfrac{|\mathcal{C}_\epsilon|}{\delta}} + \epsilon \left\|\hat{\theta} - \theta^*\right\|_V .\end{aligned}$$

(Middle step: Cauchy–Schwarz, $\langle V^{1/2}w, v\rangle \le \left\|V^{1/2}w\right\|_{V^{-1}} \|v\|_V = \|w\|_2 \|v\|_V$.) Rearrange, and take $\epsilon = 1/2$:

Fixed-design confidence ellipsoid (Lattimore–Szepesvári, eq. 20.3)

With probability at least $1 - \delta$:

$$\left\|\hat{\theta} - \theta^*\right\|_V \le \frac{1}{1-\epsilon}\sqrt{2\log\frac{|\mathcal{C}_\epsilon|}{\delta}} \le 2\sqrt{2d\log 6 + 2\log\tfrac{1}{\delta}} =: \beta_{\text{fixed}}.$$

Radius of order $\sqrt{d}$, and every direction at once:

$$|\langle z, \hat{\theta} - \theta^*\rangle| \le \|z\|_{V^{-1}} \left\|\hat{\theta} - \theta^*\right\|_V \le \beta_{\text{fixed}} \|z\|_{V^{-1}}.$$

The same shape for adaptive data (proof next lecture)

Now let x_t depend on the past. Let $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$, $\hat{\theta}_t = V_t^{-1} \sum_{s \le t} x_s y_s$ (ridge), $\|x_t\|_2 \le L$, $\|\theta^*\|_2 \le S$, 1-sub-Gaussian noise.

Theorem (Abbasi-Yadkori, Pál & Szepesvári 2011; notes Prop. 1; Lattimore–Szepesvári Thm. 20.5)

With probability at least $1 - \delta$, *simultaneously for all* $t \ge 0$,

$$\left\| \hat{\theta}_t - \theta^* \right\|_{V_t} \le \beta_t := \sqrt{\lambda}\, S + \sqrt{2 \log \tfrac{1}{\delta} + \log \tfrac{\det V_t}{\lambda^d}}$$
$$\le \sqrt{\lambda}\, S + \sqrt{2 \log \tfrac{1}{\delta} + d \log \left(1 + \tfrac{t L^2}{d \lambda}\right)}.$$

The same shape for adaptive data (proof next lecture)

Now let x_t depend on the past. Let $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$, $\hat\theta_t = V_t^{-1} \sum_{s \le t} x_s y_s$ (ridge), $\|x_t\|_2 \le L$, $\|\theta^*\|_2 \le S$, 1-sub-Gaussian noise.

Theorem (Abbasi-Yadkori, Pál & Szepesvári 2011; notes Prop. 1; Lattimore–Szepesvári Thm. 20.5)

With probability at least $1 - \delta$, *simultaneously for all* $t \ge 0$,

$$\left\| \hat\theta_t - \theta^* \right\|_{V_t} \le \beta_t := \sqrt{\lambda}\, S + \sqrt{2 \log \tfrac{1}{\delta} + \log \tfrac{\det V_t}{\lambda^d}}$$
$$\le \sqrt{\lambda}\, S + \sqrt{2 \log \tfrac{1}{\delta} + d \log \left(1 + \tfrac{t L^2}{d \lambda}\right)}.$$

- Same ellipsoid, radius $\approx \sqrt{d \log t}$ instead of $\sqrt{d}$: the $\log t$ is the price of adaptivity *and* of holding for all t at once; $\sqrt{\lambda} S$ is the ridge bias.
- No union bound over arms: the statement is about θ^*, so it covers infinitely many arms, or arm sets that change every round.
- **Next lecture:** the proof. Covering breaks when X depends on the noise; a supermartingale and the method of mixtures replace it.

From the ellipsoid to an index: optimism

Let $\mathcal{C}_t = \{\theta : \left\|\theta - \hat{\theta}_t\right\|_{V_t} \leq \beta_t\}$ and suppose $\theta^* \in \mathcal{C}_t$. Optimism says: play the arm that looks best under the most favorable θ in $\mathcal{C}_t$. For a fixed arm x that is a linear function maximized over an ellipsoid:

$$\max_{\theta \in \mathcal{C}_t} \langle x, \theta \rangle = \langle x, \hat{\theta}_t \rangle + \max_{\|v\|_{V_t} \leq \beta_t} \langle x, v \rangle =$$

From the ellipsoid to an index: optimism

Let $\mathcal{C}_t = \{\theta : \left\|\theta - \hat{\theta}_t\right\|_{V_t} \le \beta_t\}$ and suppose $\theta^* \in \mathcal{C}_t$. Optimism says: play the arm that looks best under the most favorable θ in $\mathcal{C}_t$. For a fixed arm x that is a linear function maximized over an ellipsoid:

$$\max_{\theta \in \mathcal{C}_t} \langle x, \theta \rangle = \langle x, \hat{\theta}_t \rangle + \max_{\|v\|_{V_t} \le \beta_t} \langle x, v \rangle = \langle x, \hat{\theta}_t \rangle + \beta_t \|x\|_{V_t^{-1}} .$$

(Cauchy–Schwarz, $\langle x, v \rangle = \langle V_t^{-1/2} x, V_t^{1/2} v \rangle \le \|x\|_{V_t^{-1}} \|v\|_{V_t}$, equality at $v = \beta_t V_t^{-1} x / \|x\|_{V_t^{-1}}$.) Mean plus width, as in UCB, with $\|x\|_{V_t^{-1}}$ in the role of $1/\sqrt{T_i(t)}$.

From the ellipsoid to an index: optimism

Let $\mathcal{C}_t = \{\theta : \left\|\theta - \hat{\theta}_t\right\|_{V_t} \le \beta_t\}$ and suppose $\theta^* \in \mathcal{C}_t$. Optimism says: play the arm that looks best under the most favorable θ in $\mathcal{C}_t$. For a fixed arm x that is a linear function maximized over an ellipsoid:

$$\max_{\theta \in \mathcal{C}_t} \langle x, \theta \rangle = \langle x, \hat{\theta}_t \rangle + \max_{\|v\|_{V_t} \le \beta_t} \langle x, v \rangle = \langle x, \hat{\theta}_t \rangle + \beta_t \|x\|_{V_t^{-1}} .$$

(Cauchy–Schwarz, $\langle x, v \rangle = \langle V_t^{-1/2} x, V_t^{1/2} v \rangle \le \|x\|_{V_t^{-1}} \|v\|_{V_t}$, equality at $v = \beta_t V_t^{-1} x / \|x\|_{V_t^{-1}}$.) Mean plus width, as in UCB, with $\|x\|_{V_t^{-1}}$ in the role of $1/\sqrt{T_i(t)}$.

LinUCB / OFUL (Abbasi-Yadkori et al. 2011; Lattimore–Szepesvári Ch. 19)

Input: $\lambda > 0$, δ, and the radii β_t above. **Initialize** $V_0 = \lambda I$, $\hat{\theta}_0 = 0$.
for $t = 1, 2, \ldots, T$: play $x_t = \arg\max_{x \in \mathcal{X}_t} \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$; observe y_t;

$$V_t = V_{t-1} + x_t x_t^\top, \quad \hat{\theta}_t = V_t^{-1} \textstyle\sum_{s \le t} x_s y_s .$$

$\mathcal{X}_t$ may change with t and may be infinite; for a finite set the argmax is a scan.

Regret, step 1: one round costs one width

Assume $|\langle x, \theta^* \rangle| \le 1$ for all arms; work on the good event $\theta^* \in \mathcal{C}_t$ for all t. Let $x_t^\star = \arg\max_{x \in \mathcal{X}_t} \langle x, \theta^* \rangle$, $r_t = \langle x_t^\star - x_t, \theta^* \rangle$, and $\mathrm{UCB}_t(x) = \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$. The UCB argument from Lecture 7, word for word:

$$r_t = \langle x_t^\star, \theta^* \rangle - \langle x_t, \theta^* \rangle \le$$
$$=$$

Regret, step 1: one round costs one width

Assume $|\langle x, \theta^* \rangle| \le 1$ for all arms; work on the good event $\theta^* \in \mathcal{C}_t$ for all t. Let $x_t^\star = \arg\max_{x \in \mathcal{X}_t} \langle x, \theta^* \rangle$, $r_t = \langle x_t^\star - x_t, \theta^* \rangle$, and $\mathrm{UCB}_t(x) = \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$. The UCB argument from Lecture 7, word for word:

$$
\begin{aligned}
r_t = \langle x_t^\star, \theta^* \rangle - \langle x_t, \theta^* \rangle \;&\le\; \mathrm{UCB}_t(x_t^\star) - \langle x_t, \theta^* \rangle \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \\
&= \langle x_t, \hat{\theta}_{t-1} - \theta^* \rangle + \beta_{t-1} \|x_t\|_{V_{t-1}^{-1}} \;\le\; 2\beta_{t-1} \|x_t\|_{V_{t-1}^{-1}} .
\end{aligned}
$$

The three steps: optimism ($\theta^* \in \mathcal{C}_{t-1}$, so the true value is below the UCB), the choice of x_t (largest UCB), and Cauchy–Schwarz with the ellipsoid.

Regret, step 1: one round costs one width

Assume $|\langle x, \theta^* \rangle| \le 1$ for all arms; work on the good event $\theta^* \in \mathcal{C}_t$ for all t. Let $x_t^\star = \arg\max_{x \in \mathcal{X}_t} \langle x, \theta^* \rangle$, $r_t = \langle x_t^\star - x_t, \theta^* \rangle$, and $\mathrm{UCB}_t(x) = \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$. The UCB argument from Lecture 7, word for word:

$$
\begin{aligned}
r_t = \langle x_t^\star, \theta^* \rangle - \langle x_t, \theta^* \rangle \;&\le\; \mathrm{UCB}_t(x_t^\star) - \langle x_t, \theta^* \rangle \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \\
&= \langle x_t, \hat{\theta}_{t-1} - \theta^* \rangle + \beta_{t-1} \|x_t\|_{V_{t-1}^{-1}} \;\le\; 2\beta_{t-1} \|x_t\|_{V_{t-1}^{-1}} .
\end{aligned}
$$

The three steps: optimism ($\theta^* \in \mathcal{C}_{t-1}$, so the true value is below the UCB), the choice of x_t (largest UCB), and Cauchy–Schwarz with the ellipsoid. Also $r_t \le 2$ trivially, so $r_t \le 2 \min\{1,\ \beta_{t-1} \|x_t\|_{V_{t-1}^{-1}}\} \le 2\beta_T \min\{1, \|x_t\|_{V_{t-1}^{-1}}\}$ since β_t is nondecreasing and $\beta_T \ge 1$.

What is left

In the multi-armed case the widths were $1/\sqrt{T_i(t)}$, summed arm by arm. Here the widths $\|x_t\|_{V_{t-1}^{-1}}$ shrink in a coupled way as V_t grows in the directions played. Needed: $\sum_t \min\{1, \|x_t\|^2_{V_{t-1}^{-1}}\}$ is small for *any* sequence $x_1, \ldots, x_T$.

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

For any $x_1, \dots, x_T$ with $\|x_t\|_2 \le L$ and $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$:
$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le 2 \log \frac{\det V_T}{\det V_0} \le 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$.

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

For any $x_1, \dots, x_T$ with $\|x_t\|_2 \le L$ and $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$:
$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le 2 \log \frac{\det V_T}{\det V_0} \le 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$.

Proof. Each pull grows $\det V$ by exactly the factor $1 + (\text{squared width})$:

$$\det V_t = \det\left(V_{t-1}^{1/2}\left(I + V_{t-1}^{-1/2} x_t x_t^\top V_{t-1}^{-1/2}\right) V_{t-1}^{1/2}\right) =$$

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

For any $x_1, \dots, x_T$ with $\|x_t\|_2 \le L$ and $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$:
$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le 2 \log \frac{\det V_T}{\det V_0} \le 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$.

Proof. Each pull grows $\det V$ by exactly the factor $1 + (\text{squared width})$:

$$\det V_t = \det\left(V_{t-1}^{1/2}\left(I + V_{t-1}^{-1/2} x_t x_t^\top V_{t-1}^{-1/2}\right) V_{t-1}^{1/2}\right) = \det V_{t-1}\left(1 + \|x_t\|^2_{V_{t-1}^{-1}}\right),$$

since $I + uu^\top$ has eigenvalues $1 + \|u\|_2^2$ (once) and 1 ($d-1$ times). Then $\min\{1, u\} \le 2\log(1+u)$ for $u \ge 0$, and telescope:

$$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le$$

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

For any $x_1, \dots, x_T$ with $\|x_t\|_2 \le L$ and $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$:
$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le 2 \log \frac{\det V_T}{\det V_0} \le 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$.

Proof. Each pull grows $\det V$ by exactly the factor $1 +$ (squared width):

$$\det V_t = \det\left(V_{t-1}^{1/2}\left(I + V_{t-1}^{-1/2} x_t x_t^\top V_{t-1}^{-1/2}\right) V_{t-1}^{1/2}\right) = \det V_{t-1}\left(1 + \|x_t\|^2_{V_{t-1}^{-1}}\right),$$

since $I + uu^\top$ has eigenvalues $1 + \|u\|_2^2$ (once) and 1 ($d-1$ times). Then $\min\{1, u\} \le 2\log(1+u)$ for $u \ge 0$, and telescope:

$$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \;\le\; 2 \sum_{t=1}^{T} \log \frac{\det V_t}{\det V_{t-1}} = 2 \log \frac{\det V_T}{\det V_0}.$$

Finally $\det V_T \le (\mathrm{tr} V_T / d)^d \le (\lambda + TL^2/d)^d$ (AM–GM on the eigenvalues) and $\det V_0 = \lambda^d$. ∎

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

For any $x_1, \dots, x_T$ with $\|x_t\|_2 \le L$ and $V_t = \lambda I + \sum_{s \le t} x_s x_s^\top$:
$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \le 2 \log \frac{\det V_T}{\det V_0} \le 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$.

Proof. Each pull grows $\det V$ by exactly the factor $1 +$ (squared width):

$$\det V_t = \det\left(V_{t-1}^{1/2}\left(I + V_{t-1}^{-1/2} x_t x_t^\top V_{t-1}^{-1/2}\right) V_{t-1}^{1/2}\right) = \det V_{t-1}\left(1 + \|x_t\|^2_{V_{t-1}^{-1}}\right),$$

since $I + uu^\top$ has eigenvalues $1 + \|u\|_2^2$ (once) and 1 ($d-1$ times). Then $\min\{1, u\} \le 2\log(1+u)$ for $u \ge 0$, and telescope:

$$\sum_{t=1}^{T} \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \;\le\; 2 \sum_{t=1}^{T} \log \frac{\det V_t}{\det V_{t-1}} = 2 \log \frac{\det V_T}{\det V_0}.$$

Finally $\det V_T \le (\mathrm{tr} V_T / d)^d \le (\lambda + TL^2/d)^d$ (AM–GM on the eigenvalues) and $\det V_0 = \lambda^d$. ∎

Read it: arms with width ≥ 1 can be played only $\approx d \log T$ times, since each inflates $\det V_t$ by a constant factor and $\det V_T$ is polynomial in T.

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

Under the assumptions above, with probability at least $1-\delta$, LinUCB satisfies, for all T,

$$R_T = \sum_{t=1}^{T} r_t \;\le\; 2\beta_T \sqrt{2dT \log\left(1 + \tfrac{TL^2}{d\lambda}\right)} \;=\; \widetilde{O}\big(d\sqrt{T}\big).$$

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

Under the assumptions above, with probability at least $1-\delta$, LinUCB satisfies, for all T,

$$R_T = \sum_{t=1}^{T} r_t \;\le\; 2\beta_T \sqrt{2dT \log\left(1 + \tfrac{TL^2}{d\lambda}\right)} \;=\; \widetilde{O}\big(d\sqrt{T}\big).$$

Proof. Step 1 gave $r_t \le 2\beta_T \min\{1, \|x_t\|_{V_{t-1}^{-1}}\}$ on the good event. Sum, then Cauchy–Schwarz, then step 2:

$$R_T \;\le\; 2\beta_T \sum_{t=1}^{T} \min\{1, \|x_t\|_{V_{t-1}^{-1}}\} \;\le$$

$$\le$$

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

Under the assumptions above, with probability at least $1-\delta$, LinUCB satisfies, for all T,

$$R_T = \sum_{t=1}^{T} r_t \;\le\; 2\beta_T \sqrt{2dT \log\left(1 + \tfrac{TL^2}{d\lambda}\right)} \;=\; \widetilde{O}\big(d\sqrt{T}\big).$$

Proof. Step 1 gave $r_t \le 2\beta_T \min\{1, \|x_t\|_{V_{t-1}^{-1}}\}$ on the good event. Sum, then Cauchy–Schwarz, then step 2:

$$\begin{aligned} R_T \;\le\; 2\beta_T \sum_{t=1}^{T} \min\{1, \|x_t\|_{V_{t-1}^{-1}}\} \;&\le\; 2\beta_T \sqrt{T \sum_{t=1}^{T} \min\{1, \|x_t\|^2_{V_{t-1}^{-1}}\}} \\ &\le\; 2\beta_T \sqrt{2dT \log\big(1 + TL^2/(d\lambda)\big)}. \end{aligned}$$

■

With $\beta_T \approx \sqrt{d \log T}$ this is $R_T \lesssim d\sqrt{T}\log T$: one $\sqrt{d}$ from the radius of the ellipsoid, one from the potential lemma. Nothing used that $\mathcal{X}$ is finite or fixed.

Linear Thompson sampling

The Bayesian view from Lecture 7 carries over. With a Gaussian prior $\theta^* \sim \mathcal{N}(0, \lambda^{-1} I)$ and $\mathcal{N}(0,1)$ noise, the posterior after t rounds is

$$\theta^* \mid \text{data} \;\sim\; \mathcal{N}\big(\hat{\theta}_t,\; V_t^{-1}\big), \qquad V_t = \lambda I + \sum_{s \le t} x_s x_s^\top, \quad \hat{\theta}_t = V_t^{-1} \sum_{s \le t} x_s y_s,$$

the same center and shape as the confidence ellipsoid.

Linear Thompson sampling (Agrawal & Goyal 2013)

Input: $\lambda > 0$, inflation $v \ge 1$. **Initialize** $V_0 = \lambda I$, $\hat{\theta}_0 = 0$.
for $t = 1, 2, \ldots, T$: sample $\tilde{\theta}_t \sim \mathcal{N}\big(\hat{\theta}_{t-1},\; v^2 V_{t-1}^{-1}\big)$; play $x_t = \arg\max_{x \in \mathcal{X}_t} \langle x, \tilde{\theta}_t \rangle$; observe y_t; update $V_t, \hat{\theta}_t$.

Linear Thompson sampling

The Bayesian view from Lecture 7 carries over. With a Gaussian prior $\theta^* \sim \mathcal{N}(0, \lambda^{-1} I)$ and $\mathcal{N}(0,1)$ noise, the posterior after t rounds is

$$\theta^* \mid \text{data} \;\sim\; \mathcal{N}\big(\hat{\theta}_t,\; V_t^{-1}\big), \qquad V_t = \lambda I + \sum_{s \le t} x_s x_s^\top, \quad \hat{\theta}_t = V_t^{-1} \sum_{s \le t} x_s y_s,$$

the same center and shape as the confidence ellipsoid.

Linear Thompson sampling (Agrawal & Goyal 2013)

Input: $\lambda > 0$, inflation $v \ge 1$. **Initialize** $V_0 = \lambda I$, $\hat{\theta}_0 = 0$.
for $t = 1, 2, \ldots, T$: sample $\tilde{\theta}_t \sim \mathcal{N}\big(\hat{\theta}_{t-1},\; v^2 V_{t-1}^{-1}\big)$; play $x_t = \arg\max_{x \in \mathcal{X}_t} \langle x, \tilde{\theta}_t \rangle$; observe y_t; update $V_t, \hat{\theta}_t$.

- LinUCB evaluates $\max_{\theta \in \mathcal{C}_{t-1}} \langle x, \theta \rangle$ for every arm; Thompson draws *one* $\tilde{\theta}_t$ and plays greedily against it: one linear maximization per round, which matters when $\mathcal{X}_t$ is large or combinatorial.
- Rarely-sampled directions have large posterior variance, so $\tilde{\theta}_t$ occasionally points along them: exploration with no width to compute. $v = 1$ is exactly "sample from the posterior"; the inflation $v \approx \beta_t$ is what the frequentist analysis needs.

What is known about linear Thompson sampling

- **Bayesian regret** (averaged over the prior): $\widetilde{O}(d\sqrt{T})$, matching LinUCB (Russo & Van Roy 2014). The trick: under the prior, $x^\star$ and x_t have the *same* conditional distribution given the past, so regret can be charged to confidence widths as in step 1 of the LinUCB proof.
- **Frequentist regret** (fixed θ^*): $\widetilde{O}(d^{3/2}\sqrt{T})$ with inflation $v \approx \sqrt{d \log T}$ (Agrawal & Goyal 2013; Abeille & Lazaric 2017). The proof shows the sample is optimistic, $\langle x^\star, \tilde{\theta}_t \rangle \geq \langle x^\star, \theta^* \rangle$, with constant probability each round, then runs the LinUCB argument with an extra $1/p$ and the $\sqrt{d}$ from the inflation.
- **Is the extra $\sqrt{d}$ real?** Sometimes: without inflation, linear Thompson sampling can have linear regret on carefully built instances (Hamidi & Bayati 2020).
- **In practice** it is usually the best of the three, and the inflation is rarely used: the posterior alone explores enough on typical instances.

Three algorithms, and what is next

Linear bandits so far

	elimination	LinUCB	linear Thompson
regret	$\sqrt{dT\log(nT)}$	$d\sqrt{T}\log T$	$d^{3/2}\sqrt{T}\log T$
arm set	finite, fixed	any, may change	any, may change
data used	this round only	everything	everything
per-round cost	a design	mean + width, argmax	one linear argmax
confidence tool	fixed design, union bound	ellipsoid, adaptive	posterior, inflated

- For finite $\mathcal{X}$ elimination wins by $\sqrt{d}$ ($\sqrt{dT}$ is optimal up to logs); LinUCB's extra $\sqrt{d}$ buys every direction with no union bound; Thompson trades another $\sqrt{d}$ of theory for speed, and usually wins in experiments.
- Optimism is not always right: on the Soare instance neither LinUCB nor Thompson pulls the informative but bad arm e_2, and both need α^{-4} samples to find $x^\star$ against the optimal α^{-2}.
- **Next lecture:** prove the ellipsoid for adaptive data with a supermartingale, a maximal inequality for all t, and the method of mixtures.