CSE 541: Interactive Learning

Lecture 9: Linear bandits I: least squares, experimental design, elimination

Kevin Jamieson

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

Wednesday, October 28, 2026

Last time, and where we are going

Multi-armed bandits: the complete picture

	upper bound	lower bound
regret, minimax	$\sqrt{nT \log T}$ (elimination, UCB, TS) $\sqrt{nT \log n}$ (EXP3, even adversarial)	$\frac{1}{27}\sqrt{(n-1)T}$
regret, instance	$\sum_i 2 \log T / \Delta_i$ (UCB, TS)	$\sum_i 2 \log T / \Delta_i$ (consistent)
identification	$\sum_i \Delta_i^{-2} \log(n \log \Delta_i^{-2} / \delta)$ (elimination)	$\sum_i 2\Delta_i^{-2} \log \frac{1}{4\delta}$ (δ-PAC)

Every bound has an n in it, because pulling arm i says nothing about arm j. With a million products or a continuum of doses, that is fatal.

Last time, and where we are going

Multi-armed bandits: the complete picture

	upper bound	lower bound
regret, minimax	$\sqrt{nT \log T}$ (elimination, UCB, TS) $\sqrt{nT \log n}$ (EXP3, even adversarial)	$\frac{1}{27}\sqrt{(n-1)T}$
regret, instance	$\sum_i 2 \log T / \Delta_i$ (UCB, TS)	$\sum_i 2 \log T / \Delta_i$ (consistent)
identification	$\sum_i \Delta_i^{-2} \log(n \log \Delta_i^{-2} / \delta)$ (elimination)	$\sum_i 2\Delta_i^{-2} \log \frac{1}{4\delta}$ (δ-PAC)

Every bound has an n in it, because pulling arm i says nothing about arm j. With a million products or a continuum of doses, that is fatal.

Today: arms with features. Arm i is a vector $x_i \in \mathbb{R}^d$ and its mean reward is linear in x_i. Now pulling one arm teaches you about all of them, and d should replace n. The three ingredients:

1. least squares, and how confident it is in a given *direction*;
2. experimental design: which arms to pull so that you are confident in *every* direction (Kiefer–Wolfowitz);
3. successive elimination, with the design replacing "pull every survivor equally".

Linear bandits

Least squares, one direction at a time

The stochastic linear bandit

Protocol (notes §11.1)

Known: a finite set of arms $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$ that spans $\mathbb{R}^d$. **Unknown:** $\theta^* \in \mathbb{R}^d$.
for $t = 1, 2, \dots, T$: player pulls $x_t \in \mathcal{X}$ and observes $y_t = \langle x_t, \theta^* \rangle + \eta_t$, $\eta_t \sim \mathcal{N}(0, 1)$ independent.

Let $x^\star = \arg\max_{x \in \mathcal{X}} \langle x, \theta^* \rangle$ and $\Delta_x = \langle x^\star - x, \theta^* \rangle$. The two goals are as before: regret $R_T = \mathbb{E}\big[\sum_{t \le T} \langle x^\star - x_t, \theta^* \rangle\big] = \sum_x \Delta_x \mathbb{E}[T_x]$, or finding $x^\star$ w.p. $\ge 1 - \delta$ with few pulls.

- $\mathcal{X} = \{e_1, \dots, e_d\}$ recovers the multi-armed bandit with $n = d$. Everything today must reduce to what we know in that case.
- Examples: drug candidates described by chemical descriptors; songs described by embeddings; a dose $x \in [0, 1]$ with features $(1, x, x^2, \dots)$.
- The new phenomenon: an observation at x_t is information about $\langle x, \theta^* \rangle$ for *every* x. How much? That is a question about least squares.

Least squares

Suppose we have pulled $x_1, \ldots, x_\tau$ (chosen *in advance*) and observed $y_t = \langle x_t, \theta^* \rangle + \eta_t$. Stack them: $X \in \mathbb{R}^{\tau \times d}$, $Y = X\theta^* + \eta$, $\eta \sim \mathcal{N}(0, I_\tau)$. The natural estimate minimizes the squared residuals,

$$\hat{\theta} \;=\; \operatorname*{arg\,min}_{\theta \in \mathbb{R}^d} \sum_{t=1}^{\tau} \big(y_t - \langle x_t, \theta \rangle\big)^2 \;=\; \operatorname*{arg\,min}_{\theta} \; \|Y - X\theta\|_2^2 \,.$$

Least squares

Suppose we have pulled $x_1, \ldots, x_\tau$ (chosen *in advance*) and observed $y_t = \langle x_t, \theta^* \rangle + \eta_t$. Stack them: $X \in \mathbb{R}^{\tau \times d}$, $Y = X\theta^* + \eta$, $\eta \sim \mathcal{N}(0, I_\tau)$. The natural estimate minimizes the squared residuals,

$$\hat{\theta} \;=\; \arg\min_{\theta \in \mathbb{R}^d} \sum_{t=1}^{\tau} \big(y_t - \langle x_t, \theta \rangle\big)^2 \;=\; \arg\min_{\theta} \; \|Y - X\theta\|_2^2 \,.$$

Setting the gradient to zero (the *normal equations*) and solving, assuming $X^\top X$ is invertible:

$$\begin{aligned} 0 = \nabla_\theta \, \|Y - X\theta\|_2^2 &= \\ \Longrightarrow \qquad\qquad\qquad\quad &= \end{aligned}$$

Least squares

Suppose we have pulled $x_1, \dots, x_\tau$ (chosen *in advance*) and observed $y_t = \langle x_t, \theta^* \rangle + \eta_t$. Stack them: $X \in \mathbb{R}^{\tau \times d}$, $Y = X\theta^* + \eta$, $\eta \sim \mathcal{N}(0, I_\tau)$. The natural estimate minimizes the squared residuals,

$$\hat{\theta} \;=\; \arg\min_{\theta \in \mathbb{R}^d} \sum_{t=1}^{\tau} \big(y_t - \langle x_t, \theta \rangle\big)^2 \;=\; \arg\min_{\theta} \; \|Y - X\theta\|_2^2 \,.$$

Setting the gradient to zero (the *normal equations*) and solving, assuming $X^\top X$ is invertible:

$$\begin{aligned} 0 = \nabla_\theta \|Y - X\theta\|_2^2 &= -2X^\top (Y - X\theta) \\ \implies \qquad X^\top X\, \hat{\theta} &= X^\top Y, \qquad \hat{\theta} = (X^\top X)^{-1} X^\top Y. \end{aligned}$$

Least squares

Suppose we have pulled $x_1, \ldots, x_\tau$ (chosen *in advance*) and observed $y_t = \langle x_t, \theta^* \rangle + \eta_t$. Stack them: $X \in \mathbb{R}^{\tau \times d}$, $Y = X\theta^* + \eta$, $\eta \sim \mathcal{N}(0, I_\tau)$. The natural estimate minimizes the squared residuals,

$$\hat{\theta} \;=\; \arg\min_{\theta \in \mathbb{R}^d} \sum_{t=1}^{\tau} \big(y_t - \langle x_t, \theta \rangle\big)^2 \;=\; \arg\min_{\theta} \; \|Y - X\theta\|_2^2 \,.$$

Setting the gradient to zero (the *normal equations*) and solving, assuming $X^\top X$ is invertible:

$$\begin{aligned} 0 = \nabla_\theta \|Y - X\theta\|_2^2 &= -2X^\top (Y - X\theta) \\ \implies \qquad X^\top X\, \hat{\theta} &= X^\top Y, \qquad \hat{\theta} = (X^\top X)^{-1} X^\top Y. \end{aligned}$$

Now substitute what Y actually is:

$$\hat{\theta} = (X^\top X)^{-1} X^\top (X\theta^* + \eta) =$$

Least squares

Suppose we have pulled $x_1, \ldots, x_\tau$ (chosen *in advance*) and observed $y_t = \langle x_t, \theta^* \rangle + \eta_t$. Stack them: $X \in \mathbb{R}^{\tau \times d}$, $Y = X\theta^* + \eta$, $\eta \sim \mathcal{N}(0, I_\tau)$. The natural estimate minimizes the squared residuals,

$$\hat{\theta} \;=\; \operatorname*{arg\,min}_{\theta \in \mathbb{R}^d} \sum_{t=1}^{\tau} \big(y_t - \langle x_t, \theta \rangle\big)^2 \;=\; \operatorname*{arg\,min}_{\theta} \; \|Y - X\theta\|_2^2 \,.$$

Setting the gradient to zero (the *normal equations*) and solving, assuming $X^\top X$ is invertible:

$$\begin{aligned} 0 = \nabla_\theta \|Y - X\theta\|_2^2 &= -2X^\top (Y - X\theta) \\ \implies \qquad X^\top X\, \hat{\theta} &= X^\top Y, \qquad \hat{\theta} = (X^\top X)^{-1} X^\top Y. \end{aligned}$$

Now substitute what Y actually is:

$$\hat{\theta} = (X^\top X)^{-1} X^\top (X\theta^* + \eta) = \theta^* + (X^\top X)^{-1} X^\top \eta.$$

The estimate is the truth plus a linear function of the noise. Everything about its accuracy is in the matrix $(X^\top X)^{-1} X^\top$, i.e. in the arms we pulled.

The error in a direction

We care about the error in a *direction* $z \in \mathbb{R}^d$: $z = x$ for "how good is arm x", $z = x - x'$ for "which of two arms is better". From the last slide,

$$\langle z, \hat{\theta} - \theta^* \rangle = z^\top (X^\top X)^{-1} X^\top \eta = \langle w, \eta \rangle, \qquad w := X(X^\top X)^{-1} z \in \mathbb{R}^\tau.$$

Here w is a fixed vector: it depends on the design and on z, not on the noise.

The error in a direction

We care about the error in a *direction* $z \in \mathbb{R}^d$: $z = x$ for "how good is arm x", $z = x - x'$ for "which of two arms is better". From the last slide,

$$\langle z, \hat{\theta} - \theta^* \rangle = z^\top (X^\top X)^{-1} X^\top \eta = \langle w, \eta \rangle, \qquad w := X(X^\top X)^{-1} z \in \mathbb{R}^\tau.$$

Here w is a fixed vector: it depends on the design and on z, not on the noise. A fixed linear combination of independent $\mathcal{N}(0,1)$'s is $\mathcal{N}(0, \|w\|_2^2)$, and

$$\|w\|_2^2 =$$

The error in a direction

We care about the error in a *direction* $z \in \mathbb{R}^d$: $z = x$ for "how good is arm x", $z = x - x'$ for "which of two arms is better". From the last slide,

$$\langle z, \hat{\theta} - \theta^* \rangle = z^\top (X^\top X)^{-1} X^\top \eta = \langle w, \eta \rangle, \qquad w := X(X^\top X)^{-1} z \in \mathbb{R}^\tau.$$

Here w is a fixed vector: it depends on the design and on z, not on the noise. A fixed linear combination of independent $\mathcal{N}(0,1)$'s is $\mathcal{N}(0, \|w\|_2^2)$, and

$$\|w\|_2^2 = z^\top (X^\top X)^{-1} X^\top X (X^\top X)^{-1} z = z^\top (X^\top X)^{-1} z =: \|z\|^2_{(X^\top X)^{-1}}.$$

The error in a direction

We care about the error in a *direction* $z \in \mathbb{R}^d$: $z = x$ for "how good is arm x", $z = x - x'$ for "which of two arms is better". From the last slide,

$$\langle z, \hat{\theta} - \theta^* \rangle = z^\top (X^\top X)^{-1} X^\top \eta = \langle w, \eta \rangle, \qquad w := X(X^\top X)^{-1} z \in \mathbb{R}^\tau.$$

Here w is a fixed vector: it depends on the design and on z, not on the noise. A fixed linear combination of independent $\mathcal{N}(0,1)$'s is $\mathcal{N}(0, \|w\|_2^2)$, and

$$\|w\|_2^2 = z^\top (X^\top X)^{-1} X^\top X (X^\top X)^{-1} z = z^\top (X^\top X)^{-1} z =: \|z\|^2_{(X^\top X)^{-1}}.$$

Fixed-design confidence interval

For any fixed z, with probability at least $1 - \delta$:
$\langle z, \hat{\theta} - \theta^* \rangle \;\le\; \|z\|_{(X^\top X)^{-1}} \sqrt{2 \log(1/\delta)}$.

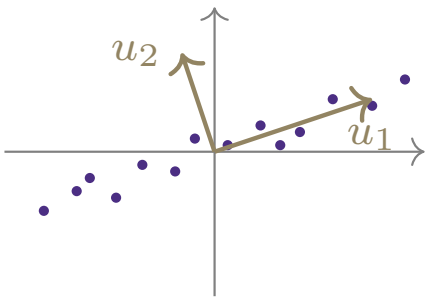
Sanity check: $\mathcal{X} = \{e_i\}$, arm i pulled τ_i times, $z = e_i$:
$X^\top X = \mathrm{diag}(\tau_1, \ldots, \tau_d)$, so $\|e_i\|^2_{(X^\top X)^{-1}} = 1/\tau_i$ and we recover $\sqrt{2\log(1/\delta)/\tau_i}$. Note also $\|e_1 - e_2\|^2_{(X^\top X)^{-1}} = 1/\tau_1 + 1/\tau_2$: comparing two arms costs the sum of their variances, as it should.

A picture: which directions are uncertain?

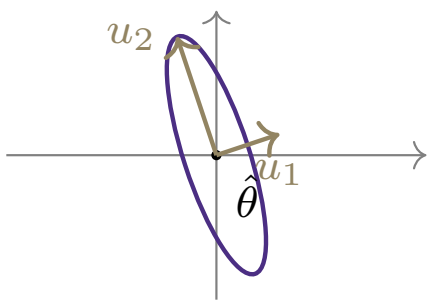
Write the SVD $X = VSU^\top$ with $U = [u_1, \dots, u_d]$ orthonormal and singular values $s_1 \geq \cdots \geq s_d > 0$. Then $X^\top X = US^2U^\top$ and

$$\|z\|^2_{(X^\top X)^{-1}} = \sum_{j=1}^{d} \frac{\langle z, u_j \rangle^2}{s_j^2},$$

and $\{\theta : \left\|\theta - \hat{\theta}\right\|^2_{X^\top X} \leq \beta\}$, which contains θ^* w.h.p. for $\beta \approx d$, is an *ellipsoid* with semi-axis $\sqrt{\beta}/s_j$ along u_j.



the pulls: many along u_1, few along u_2



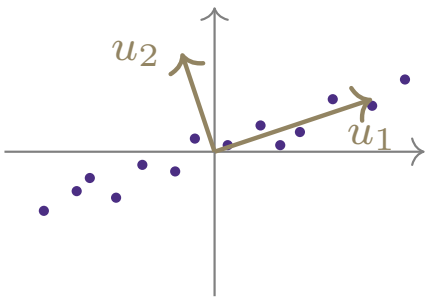
where θ^* can be: wide along u_2

A picture: which directions are uncertain?

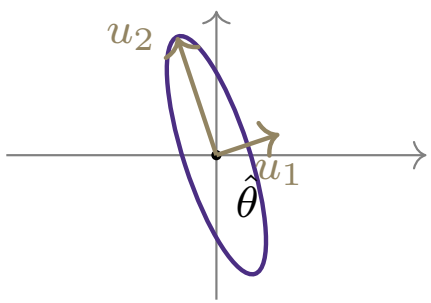
Write the SVD $X = VSU^\top$ with $U = [u_1, \dots, u_d]$ orthonormal and singular values $s_1 \geq \cdots \geq s_d > 0$. Then $X^\top X = US^2U^\top$ and

$$\|z\|^2_{(X^\top X)^{-1}} = \sum_{j=1}^{d} \frac{\langle z, u_j \rangle^2}{s_j^2},$$

and $\{\theta : \left\|\theta - \hat{\theta}\right\|^2_{X^\top X} \leq \beta\}$, which contains θ^* w.h.p. for $\beta \approx d$, is an *ellipsoid* with semi-axis $\sqrt{\beta}/s_j$ along u_j.



the pulls: many along u_1, few along u_2



where θ^* can be: wide along u_2

The directions you sampled a lot are the directions you know; an arm along u_2 has a wide interval however many pulls were made along u_1. Choosing the pulls is choosing the shape of this ellipsoid.

Sub-Gaussian noise: Chernoff, one more time

Only used: $\langle z, \hat{\theta} - \theta^* \rangle = \langle w, \eta \rangle$ with w fixed and $\eta_1, \ldots, \eta_\tau$ independent 1-sub-Gaussian. For $\mu > 0$ and any $\lambda > 0$,

$$\mathbb{P}\big(\langle w, \eta \rangle > \mu\big) \leq e^{-\lambda\mu}\, \mathbb{E}\big[e^{\lambda \langle w, \eta \rangle}\big] = e^{-\lambda\mu} \prod_{s=1}^{\tau} \mathbb{E}\big[e^{\lambda w_s \eta_s}\big]$$

$$\leq$$

$$\leq$$

Sub-Gaussian noise: Chernoff, one more time

Only used: $\langle z, \hat{\theta} - \theta^* \rangle = \langle w, \eta \rangle$ with w fixed and $\eta_1, \dots, \eta_\tau$ independent 1-sub-Gaussian. For $\mu > 0$ and any $\lambda > 0$,

$$\begin{aligned}
\mathbb{P}\big(\langle w, \eta \rangle > \mu\big) &\le e^{-\lambda\mu}\, \mathbb{E}\big[e^{\lambda \langle w, \eta \rangle}\big] = e^{-\lambda\mu} \prod_{s=1}^{\tau} \mathbb{E}\big[e^{\lambda w_s \eta_s}\big] \\
&\le e^{-\lambda\mu} \prod_{s=1}^{\tau} e^{\lambda^2 w_s^2/2} = \exp\Big(-\lambda\mu + \tfrac{\lambda^2}{2} \|w\|_2^2 \Big) \\
&\le \exp\Big(-\frac{\mu^2}{2\|w\|_2^2} \Big) \quad \text{at } \lambda = \mu / \|w\|_2^2 .
\end{aligned}$$

Setting this to δ gives $\mu = \|w\|_2 \sqrt{2\log(1/\delta)} = \|z\|_{(X^\top X)^{-1}} \sqrt{2\log(1/\delta)}$, the same interval.

Sub-Gaussian noise: Chernoff, one more time

Only used: $\langle z, \hat{\theta} - \theta^* \rangle = \langle w, \eta \rangle$ with w fixed and $\eta_1, \dots, \eta_\tau$ independent 1-sub-Gaussian. For $\mu > 0$ and any $\lambda > 0$,

$$\begin{aligned}
\mathbb{P}\big(\langle w, \eta \rangle > \mu\big) &\le e^{-\lambda\mu}\, \mathbb{E}\big[e^{\lambda \langle w, \eta \rangle}\big] = e^{-\lambda\mu} \prod_{s=1}^{\tau} \mathbb{E}\big[e^{\lambda w_s \eta_s}\big] \\
&\le e^{-\lambda\mu} \prod_{s=1}^{\tau} e^{\lambda^2 w_s^2/2} = \exp\Big(-\lambda\mu + \tfrac{\lambda^2}{2} \|w\|_2^2 \Big) \\
&\le \exp\Big(-\frac{\mu^2}{2\|w\|_2^2} \Big) \quad \text{at } \lambda = \mu / \|w\|_2^2 .
\end{aligned}$$

Setting this to δ gives $\mu = \|w\|_2 \sqrt{2\log(1/\delta)} = \|z\|_{(X^\top X)^{-1}} \sqrt{2\log(1/\delta)}$, the same interval.

Where the independence was used

The product over s needs w fixed, i.e. the design $x_1, \dots, x_\tau$ chosen before any noise is seen. A bandit algorithm chooses x_t *because of* past noise, and then this calculation is false (the error can inflate by a factor d). Today: fresh samples each round. Next lecture: the martingale fix.

Experimental design

Choosing where to measure so that every direction is well estimated

Designs, and what to optimize

The error in direction z is $\|z\|_{(X^\top X)^{-1}}$: it depends on the arms we pulled only through $X^\top X = \sum_{t \le \tau} x_t x_t^\top$. So *choose* the pulls to shape it. When τ is large it is convenient to think of a distribution $\lambda \in \triangle_{\mathcal{X}}$ over arms, a design, with

$$A(\lambda) = \sum_{x \in \mathcal{X}} \lambda_x \, x x^\top .$$

Pulling x about $\lambda_x \tau$ times gives $X^\top X \approx \tau A(\lambda)$, hence $\|z\|^2_{(X^\top X)^{-1}} \approx \frac{1}{\tau} \|z\|^2_{A(\lambda)^{-1}}$: the design sets the shape, τ the scale.

Designs, and what to optimize

The error in direction z is $\|z\|_{(X^\top X)^{-1}}$: it depends on the arms we pulled only through $X^\top X = \sum_{t \le \tau} x_t x_t^\top$. So *choose* the pulls to shape it. When τ is large it is convenient to think of a distribution $\lambda \in \triangle_{\mathcal{X}}$ over arms, a design, with

$$A(\lambda) = \sum_{x \in \mathcal{X}} \lambda_x \, x x^\top .$$

Pulling x about $\lambda_x \tau$ times gives $X^\top X \approx \tau A(\lambda)$, hence $\|z\|^2_{(X^\top X)^{-1}} \approx \frac{1}{\tau} \|z\|^2_{A(\lambda)^{-1}}$: the design sets the shape, τ the scale.

Classical criteria (notes §11.3)

name	minimize	controls
A-optimal	$\operatorname{tr} A(\lambda)^{-1}$	$\mathbb{E}\left\|\hat{\theta} - \theta^*\right\|_2^2$
E-optimal	$\max_{\|u\|_2 \le 1} \|u\|^2_{A(\lambda)^{-1}}$	worst Euclidean direction
D-optimal	$-\log\det A(\lambda)$	volume of the confidence ellipsoid
G-optimal	$\max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}}$	worst arm: $\max_x \mathbb{E}[\langle x, \hat{\theta} - \theta^* \rangle^2]$

For bandits the G criterion is the one we want: it is the squared width of the confidence interval of the worst arm. Write $f(\lambda) = \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}}$.

Kiefer–Wolfowitz: the worst arm has leverage exactly d

Theorem (Kiefer & Wolfowitz 1960; notes Lemma 19)

Let $\mathcal{X} \subset \mathbb{R}^d$ be finite and span $\mathbb{R}^d$. Then $\min_{\lambda \in \triangle_{\mathcal{X}}} f(\lambda) = \min_{\lambda} \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}} = d$, the minimizers of f coincide with the maximizers of $\log\det A(\lambda)$ (D-optimal $=$ G-optimal), and there is a minimizer supported on at most $d(d+1)/2$ arms.

No proof here (Lattimore–Szepesvári Ch. 21). Two things to take from it:

Kiefer–Wolfowitz: the worst arm has leverage exactly d

Theorem (Kiefer & Wolfowitz 1960; notes Lemma 19)

Let $\mathcal{X} \subset \mathbb{R}^d$ be finite and span $\mathbb{R}^d$. Then $\min_{\lambda \in \triangle_{\mathcal{X}}} f(\lambda) = \min_{\lambda} \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}} = d$, the minimizers of f coincide with the maximizers of $\log\det A(\lambda)$ (D-optimal = G-optimal), and there is a minimizer supported on at most $d(d+1)/2$ arms.

No proof here (Lattimore–Szepesvári Ch. 21). Two things to take from it:

- **The value d does not depend on the geometry of $\mathcal{X}$, or on n.** For $\mathcal{X} = \{e_i\}$: $A(\lambda) = \mathrm{diag}(\lambda)$, $f(\lambda) = \max_i 1/\lambda_i$, minimized by the uniform design, value d: pull every arm equally. Every arm set is as good as this one once the design is chosen well. (And no design beats d: $\sum_x \lambda_x \|x\|^2_{A(\lambda)^{-1}} = \mathrm{tr}(A(\lambda)^{-1} A(\lambda)) = d$, so the *average* leverage is d.)

Kiefer–Wolfowitz: the worst arm has leverage exactly d

Theorem (Kiefer & Wolfowitz 1960; notes Lemma 19)

Let $\mathcal{X} \subset \mathbb{R}^d$ be finite and span $\mathbb{R}^d$. Then $\min_{\lambda \in \triangle_{\mathcal{X}}} f(\lambda) = \min_\lambda \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}} = d$, the minimizers of f coincide with the maximizers of $\log\det A(\lambda)$ (D-optimal = G-optimal), and there is a minimizer supported on at most $d(d+1)/2$ arms.

No proof here (Lattimore–Szepesvári Ch. 21). Two things to take from it:

- **The value d does not depend on the geometry of $\mathcal{X}$, or on n.** For $\mathcal{X} = \{e_i\}$: $A(\lambda) = \mathrm{diag}(\lambda)$, $f(\lambda) = \max_i 1/\lambda_i$, minimized by the uniform design, value d: pull every arm equally. Every arm set is as good as this one once the design is chosen well. (And no design beats d: $\sum_x \lambda_x \|x\|^2_{A(\lambda)^{-1}} = \mathrm{tr}(A(\lambda)^{-1}A(\lambda)) = d$, so the *average* leverage is d.)
- **Where $d(d+1)/2$ comes from: Carathéodory.** $A(\lambda) = \sum_x \lambda_x x x^\top$ is a convex combination of the matrices $xx^\top$, points in the space of symmetric $d \times d$ matrices, which has dimension $d(d+1)/2$. Carathéodory: a point in the convex hull of a set in $\mathbb{R}^m$ is a convex combination of at most $m+1$ of its points. So $A(\lambda^\star)$ can be written using at most $d(d+1)/2+1$ arms (one more can be shaved off), and f depends on λ only through $A(\lambda)$. However many arms there are, a good design pulls $O(d^2)$ of them.

What a G-optimal design buys

Proposition (notes Prop. 7)

Let $\lambda^\star$ be G-optimal for $\mathcal{X}$ with $|\mathrm{supp}(\lambda^\star)| \le d(d+1)/2$. Pull each x exactly $\lceil \tau \lambda^\star_x \rceil$ times (at most $\tau + d(d+1)/2$ pulls) and compute least squares $\hat{\theta}$. Then w.p. $\ge 1-\delta$, simultaneously for all $x \in \mathcal{X}$, $\quad |\langle x, \hat{\theta} - \theta^* \rangle| \le \sqrt{2d \log(2n/\delta)/\tau}$.

What a G-optimal design buys

Proposition (notes Prop. 7)

Let $\lambda^\star$ be G-optimal for $\mathcal{X}$ with $|\mathrm{supp}(\lambda^\star)| \le d(d+1)/2$. Pull each x exactly $\lceil \tau \lambda^\star_x \rceil$ times (at most $\tau + d(d+1)/2$ pulls) and compute least squares $\hat{\theta}$. Then w.p. $\ge 1-\delta$, simultaneously for all $x \in \mathcal{X}$, $\quad |\langle x, \hat{\theta} - \theta^* \rangle| \le \sqrt{2d \log(2n/\delta)/\tau}$.

Proof. Fix x: the fixed-design interval at level $\delta/(2n)$ (union bound over n arms, two signs); then $X^\top X \succeq \tau A(\lambda^\star)$, so $(X^\top X)^{-1} \preceq \tau^{-1} A(\lambda^\star)^{-1}$; then Kiefer–Wolfowitz:

$$
\begin{aligned}
|\langle x, \hat{\theta} - \theta^* \rangle| &\le \|x\|_{(X^\top X)^{-1}} \sqrt{2 \log(2n/\delta)} \\
&\le
\end{aligned}
$$

What a G-optimal design buys

Proposition (notes Prop. 7)

Let $\lambda^\star$ be G-optimal for $\mathcal{X}$ with $|\mathrm{supp}(\lambda^\star)| \le d(d+1)/2$. Pull each x exactly $\lceil \tau \lambda^\star_x \rceil$ times (at most $\tau + d(d+1)/2$ pulls) and compute least squares $\hat\theta$. Then w.p. $\ge 1-\delta$, simultaneously for all $x \in \mathcal{X}$, $\quad |\langle x, \hat\theta - \theta^* \rangle| \le \sqrt{2d \log(2n/\delta)/\tau}$.

Proof. Fix x: the fixed-design interval at level $\delta/(2n)$ (union bound over n arms, two signs); then $X^\top X \succeq \tau A(\lambda^\star)$, so $(X^\top X)^{-1} \preceq \tau^{-1} A(\lambda^\star)^{-1}$; then Kiefer–Wolfowitz:

$$
\begin{aligned}
|\langle x, \hat\theta - \theta^* \rangle| &\le \|x\|_{(X^\top X)^{-1}} \sqrt{2\log(2n/\delta)} \\
&\le \frac{\|x\|_{A(\lambda^\star)^{-1}}}{\sqrt{\tau}} \sqrt{2\log(2n/\delta)} \le \sqrt{\frac{2d\log(2n/\delta)}{\tau}}. \qquad \blacksquare
\end{aligned}
$$

What a G-optimal design buys

Proposition (notes Prop. 7)

Let $\lambda^\star$ be G-optimal for $\mathcal{X}$ with $|\mathrm{supp}(\lambda^\star)| \le d(d+1)/2$. Pull each x exactly $\lceil \tau \lambda^\star_x \rceil$ times (at most $\tau + d(d+1)/2$ pulls) and compute least squares $\hat\theta$. Then w.p. $\ge 1-\delta$, simultaneously for all $x \in \mathcal{X}$, $\quad |\langle x, \hat\theta - \theta^* \rangle| \le \sqrt{2d \log(2n/\delta)/\tau}$.

Proof. Fix x: the fixed-design interval at level $\delta/(2n)$ (union bound over n arms, two signs); then $X^\top X \succeq \tau A(\lambda^\star)$, so $(X^\top X)^{-1} \preceq \tau^{-1} A(\lambda^\star)^{-1}$; then Kiefer–Wolfowitz:

$$
\begin{aligned}
|\langle x, \hat\theta - \theta^* \rangle| &\le \|x\|_{(X^\top X)^{-1}} \sqrt{2 \log(2n/\delta)} \\
&\le \frac{\|x\|_{A(\lambda^\star)^{-1}}}{\sqrt{\tau}} \sqrt{2 \log(2n/\delta)} \le \sqrt{\frac{2d \log(2n/\delta)}{\tau}}. \qquad \blacksquare
\end{aligned}
$$

Compare with the multi-armed bandit

Pulling each of n arms τ/n times gives width $\sqrt{2n \log(2n/\delta)/\tau}$. With features, n becomes d and the number of arms enters only through $\log n$. That is the entire gain of linear bandits, a fact about least squares, not about any algorithm.

Computing a design, 1: an algorithm you already know

By Kiefer–Wolfowitz it suffices to solve the D-optimal problem $\max_{\lambda \in \triangle_{\mathcal{X}}} g(\lambda)$, $g(\lambda) = \log\det A(\lambda)$: a *concave* function over the *simplex*, with gradient $[\nabla g(\lambda)]_x = \mathrm{tr}\big(A(\lambda)^{-1} x x^\top\big) = \|x\|^2_{A(\lambda)^{-1}}$, the leverage of arm x.

Computing a design, 1: an algorithm you already know

By Kiefer–Wolfowitz it suffices to solve the D-optimal problem $\max_{\lambda \in \triangle_{\mathcal{X}}} g(\lambda)$, $g(\lambda) = \log\det A(\lambda)$: a *concave* function over the *simplex*, with gradient $[\nabla g(\lambda)]_x = \mathrm{tr}\big(A(\lambda)^{-1} x x^\top\big) = \|x\|^2_{A(\lambda)^{-1}}$, the leverage of arm x. Concave maximization over the simplex is what exponential weights does:

Exponential weights for D-optimal design

Input: $\mathcal{X}$, step size η. **Initialize** λ^1 uniform on $\mathcal{X}$.
for $k = 1, 2, \ldots$: $\quad \lambda_x^{k+1} \;\propto\; \lambda_x^k \exp\left(\eta \, \|x\|^2_{A(\lambda^k)^{-1}}\right)$ for every $x \in \mathcal{X}$.

Each step boosts the arms that are currently worst estimated; the usual analysis gives $g(\lambda^\star) - g(\bar{\lambda}^K) \lesssim \sqrt{\log(n)/K}$.

Computing a design, 1: an algorithm you already know

By Kiefer–Wolfowitz it suffices to solve the D-optimal problem $\max_{\lambda \in \triangle_{\mathcal{X}}} g(\lambda)$, $g(\lambda) = \log\det A(\lambda)$: a *concave* function over the *simplex*, with gradient $[\nabla g(\lambda)]_x = \mathrm{tr}\big(A(\lambda)^{-1} x x^\top\big) = \|x\|^2_{A(\lambda)^{-1}}$, the leverage of arm x. Concave maximization over the simplex is what exponential weights does:

Exponential weights for D-optimal design

Input: $\mathcal{X}$, step size η. **Initialize** λ^1 uniform on $\mathcal{X}$.
for $k = 1, 2, \ldots$: $\quad \lambda_x^{k+1} \;\propto\; \lambda_x^k \exp\left(\eta \, \|x\|^2_{A(\lambda^k)^{-1}}\right)$ for every $x \in \mathcal{X}$.

Each step boosts the arms that are currently worst estimated; the usual analysis gives $g(\lambda^\star) - g(\bar{\lambda}^K) \lesssim \sqrt{\log(n)/K}$.

- The output is *dense*: every arm gets some weight. Fine for sampling, but we promised $O(d^2)$ support.
- **Sparsify with Carathéodory, constructively.** If $|\mathrm{supp}(\lambda)| > d(d+1)/2 + 1$, the matrices $xx^\top$ in the support are linearly dependent: find $c \neq 0$ with $\sum_x c_x \, xx^\top = 0$, $\sum_x c_x = 0$, and set $\lambda \leftarrow \lambda - tc$ with the largest t keeping $\lambda \geq 0$. One arm leaves the support and $A(\lambda)$ is unchanged. Repeat.

Computing a design, 2: Frank–Wolfe, sparse by construction

Frank–Wolfe for the same concave program moves toward the vertex of the simplex with the largest gradient, i.e. toward the single arm with the largest leverage:

Frank–Wolfe for D/G-optimal design (notes §11.3.1)

Input: $\mathcal{X} = \{x_1, \dots, x_n\}$, initial $\lambda^1 \in \triangle_n$ (e.g. uniform on $2d$ random arms that span $\mathbb{R}^d$).
for $k = 1, 2, \dots$: $\quad I_k = \arg\max_i \|x_i\|^2_{A(\lambda^k)^{-1}}$; $\quad \lambda^{k+1} = (1-\gamma_k)\lambda^k + \gamma_k e_{I_k}$ with γ_k by line search (or $2/(k+1)$);
$\quad$ **stop** when $\max_i \|x_i\|^2_{A(\lambda^{k+1})^{-1}} \le 2d$.

Computing a design, 2: Frank–Wolfe, sparse by construction

Frank–Wolfe for the same concave program moves toward the vertex of the simplex with the largest gradient, i.e. toward the single arm with the largest leverage:

Frank–Wolfe for D/G-optimal design (notes §11.3.1)

Input: $\mathcal{X} = \{x_1, \dots, x_n\}$, initial $\lambda^1 \in \triangle_n$ (e.g. uniform on $2d$ random arms that span $\mathbb{R}^d$).
for $k = 1, 2, \dots$: $\quad I_k = \arg\max_i \|x_i\|^2_{A(\lambda^k)^{-1}}$; $\quad \lambda^{k+1} = (1-\gamma_k)\lambda^k + \gamma_k e_{I_k}$ with γ_k by line search (or $2/(k+1)$);
stop when $\max_i \|x_i\|^2_{A(\lambda^{k+1})^{-1}} \le 2d$.

- Each step adds at most one arm to the support: sparse by construction.
- **Fact** (no proof; notes §11.3.1): with line search it stops after $O(d \log d)$ steps, so a design with $f(\lambda) \le 2d$ and support $O(d \log d)$ is cheap for any n. The factor 2 costs a factor 2 in pulls and nothing else.
- The same leverage $\|x\|^2_{A(\lambda)^{-1}}$ drives both algorithms: the gradient of $\log\det$, the squared width of arm x, and above d for some arm unless the design is optimal.

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 $\nu \approx \sqrt{d \log(\cdot)/T}$.

What we have, and what is not quite right yet

- d **replaces** n. $\sqrt{dT\log(nT)}$ against $\sqrt{nT\log T}$; for $\mathcal{X}=\{e_i\}$ they coincide, and last time showed that is minimax optimal up to logs. A million arms in ten dimensions cost like ten arms.
- **Pure exploration for free, suspiciously.** The lemma identifies $x^\star$ after $\lesssim \frac{d}{\Delta^2}\log(n\log(1/\Delta)/\delta)$ pulls. But a *single* G-optimal design with $\tau \approx d\Delta^{-2}\log(n/\delta)$ pulls, no adaptivity at all, gets the same. The right adaptive rate can be much smaller; we will come back to it.
- **The gap-dependent bound is a little off.** For $\mathcal{X}=\{e_i\}$ it reads $\frac{d}{\Delta}\log T$ while UCB gets $\sum_i \frac{\log T}{\Delta_i}$: elimination spreads a whole design over all survivors, so it over-samples arms with large gaps. Optimism (LinUCB) is the natural fix, and the next lecture builds the confidence set it needs.
- **Fresh samples each round was a trick.** We threw away data to keep the design fixed; reusing it all makes the design depend on past noise, and the fixed-design interval is false. Next: a confidence ellipsoid for adaptively collected data, for all t at once, via martingales and the method of mixtures.