CSE 541: Interactive Learning

Lecture 12: Martingale confidence sets, and adversarial linear bandits

Kevin Jamieson

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

Monday, November 9, 2026

Martingales

First, the vocabulary: probability spaces, information, conditioning

Probability spaces and σ-algebras

A **probability space** is $(\Omega, \mathcal{F}, \mathbb{P})$: a set of outcomes Ω, a collection $\mathcal{F}$ of subsets of Ω called *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$.

σ-algebra

$\mathcal{F}$ is a σ-algebra if it contains Ω and is closed under complements and countable unions. $\mathbb{P}$ is countably additive on $\mathcal{F}$.

Probability spaces and σ-algebras

A **probability space** is $(\Omega, \mathcal{F}, \mathbb{P})$: a set of outcomes Ω, a collection $\mathcal{F}$ of subsets of Ω called *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$.

σ-algebra

$\mathcal{F}$ is a σ-algebra if it contains Ω and is closed under complements and countable unions. $\mathbb{P}$ is countably additive on $\mathcal{F}$.

- **Read a σ-algebra as information.** $\mathcal{F}$ is the set of yes/no questions about ω that can be answered. A smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is less information.

Probability spaces and σ-algebras

A **probability space** is $(\Omega, \mathcal{F}, \mathbb{P})$: a set of outcomes Ω, a collection $\mathcal{F}$ of subsets of Ω called *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$.

σ-algebra

$\mathcal{F}$ is a σ-algebra if it contains Ω and is closed under complements and countable unions. $\mathbb{P}$ is countably additive on $\mathcal{F}$.

- **Read a σ-algebra as information.** $\mathcal{F}$ is the set of yes/no questions about ω that can be answered. A smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is less information.
- **Bandit example.** $\Omega =$ all transcripts $(x_1, y_1, x_2, y_2, \dots)$ plus the algorithm's coins; $\mathcal{F}_t = \sigma(x_1, y_1, \dots, x_t, y_t)$ is what is answerable after round t: "was $x_3 = e_2$?" is in $\mathcal{F}_3$; "is $y_5 > 0$?" is not.

Probability spaces and σ-algebras

A **probability space** is $(\Omega, \mathcal{F}, \mathbb{P})$: a set of outcomes Ω, a collection $\mathcal{F}$ of subsets of Ω called *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$.

σ-algebra

$\mathcal{F}$ is a σ-algebra if it contains Ω and is closed under complements and countable unions. $\mathbb{P}$ is countably additive on $\mathcal{F}$.

- **Read a σ-algebra as information.** $\mathcal{F}$ is the set of yes/no questions about ω that can be answered. A smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is less information.
- **Bandit example.** $\Omega =$ all transcripts $(x_1, y_1, x_2, y_2, \dots)$ plus the algorithm's coins; $\mathcal{F}_t = \sigma(x_1, y_1, \dots, x_t, y_t)$ is what is answerable after round t: "was $x_3 = e_2$?" is in $\mathcal{F}_3$; "is $y_5 > 0$?" is not.
- **Filtration.** $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$: information only accumulates. All of today lives on a filtered probability space.

Probability spaces and σ-algebras

A **probability space** is $(\Omega, \mathcal{F}, \mathbb{P})$: a set of outcomes Ω, a collection $\mathcal{F}$ of subsets of Ω called *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$.

σ-algebra

$\mathcal{F}$ is a σ-algebra if it contains Ω and is closed under complements and countable unions. $\mathbb{P}$ is countably additive on $\mathcal{F}$.

- **Read a σ-algebra as information.** $\mathcal{F}$ is the set of yes/no questions about ω that can be answered. A smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is less information.
- **Bandit example.** $\Omega =$ all transcripts $(x_1, y_1, x_2, y_2, \dots)$ plus the algorithm's coins; $\mathcal{F}_t = \sigma(x_1, y_1, \dots, x_t, y_t)$ is what is answerable after round t: "was $x_3 = e_2$?" is in $\mathcal{F}_3$; "is $y_5 > 0$?" is not.
- **Filtration.** $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$: information only accumulates. All of today lives on a filtered probability space.

Nothing deep: bookkeeping for "what was known when", which is exactly what adaptive algorithms force us to track.

Random variables, measurability, conditional expectation

A **random variable** is a function $X : \Omega \to \mathbb{R}$ with $\{X \leq a\} \in \mathcal{F}$ for every a. It is $\mathcal{G}$-**measurable** if those events lie in the sub-σ-algebra $\mathcal{G}$: the information in $\mathcal{G}$ determines X.

Random variables, measurability, conditional expectation

A **random variable** is a function $X : \Omega \to \mathbb{R}$ with $\{X \le a\} \in \mathcal{F}$ for every a. It is $\mathcal{G}$-**measurable** if those events lie in the sub-σ-algebra $\mathcal{G}$: the information in $\mathcal{G}$ determines X.

- Z_t is **adapted** if Z_t is $\mathcal{F}_t$-measurable (known after round t), **predictable** if $\mathcal{F}_{t-1}$-measurable (decided before round t's noise). In a bandit the arm x_t is predictable; y_t and η_t are adapted.

Random variables, measurability, conditional expectation

A **random variable** is a function $X : \Omega \to \mathbb{R}$ with $\{X \le a\} \in \mathcal{F}$ for every a. It is $\mathcal{G}$-**measurable** if those events lie in the sub-σ-algebra $\mathcal{G}$: the information in $\mathcal{G}$ determines X.

- Z_t is **adapted** if Z_t is $\mathcal{F}_t$-measurable (known after round t), **predictable** if $\mathcal{F}_{t-1}$-measurable (decided before round t's noise). In a bandit the arm x_t is predictable; y_t and η_t are adapted.

Conditional expectation

$\mathbb{E}[X \mid \mathcal{G}]$: the best prediction of X from the information in $\mathcal{G}$, i.e. the $\mathcal{G}$-measurable Y with $\mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[X\mathbf{1}_A]$ for all $A \in \mathcal{G}$.

Two rules used constantly: **taking out what is known**, $\mathbb{E}[YX \mid \mathcal{G}] = Y\,\mathbb{E}[X \mid \mathcal{G}]$ for $\mathcal{G}$-measurable Y (given $\mathcal{F}_{t-1}$, x_t is a constant); and the **tower rule**, $\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}]$ for $\mathcal{G} \subseteq \mathcal{H}$, so $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_{t-1}]] = \mathbb{E}[X]$.

Random variables, measurability, conditional expectation

A **random variable** is a function $X : \Omega \to \mathbb{R}$ with $\{X \le a\} \in \mathcal{F}$ for every a. It is $\mathcal{G}$-**measurable** if those events lie in the sub-σ-algebra $\mathcal{G}$: the information in $\mathcal{G}$ determines X.

- Z_t is **adapted** if Z_t is $\mathcal{F}_t$-measurable (known after round t), **predictable** if $\mathcal{F}_{t-1}$-measurable (decided before round t's noise). In a bandit the arm x_t is predictable; y_t and η_t are adapted.

Conditional expectation

$\mathbb{E}[X \mid \mathcal{G}]$: the best prediction of X from the information in $\mathcal{G}$, i.e. the $\mathcal{G}$-measurable Y with $\mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[X\mathbf{1}_A]$ for all $A \in \mathcal{G}$.

Two rules used constantly: **taking out what is known**, $\mathbb{E}[YX \mid \mathcal{G}] = Y\,\mathbb{E}[X \mid \mathcal{G}]$ for $\mathcal{G}$-measurable Y (given $\mathcal{F}_{t-1}$, x_t is a constant); and the **tower rule**, $\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}]$ for $\mathcal{G} \subseteq \mathcal{H}$, so $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_{t-1}]] = \mathbb{E}[X]$. **Bandit noise, stated properly:** $\mathbb{E}[e^{\lambda\eta_t} \mid \mathcal{F}_{t-1}] \le e^{\lambda^2/2}$ for all λ: sub-Gaussian *given* the past; independence is not required.

Martingales

Confidence bounds that hold for all time, and for adaptive designs

Two problems, one tool

1. **For all t at once.** Chernoff bounds S_t at one fixed t; UCB needed all $t \le T$ and paid a union bound. What is the right price for every t simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \dots, y_{t-1}$: the summands $x_t \eta_t$ are not independent and the fixed-design calculation is wrong.

Two problems, one tool

1. **For all t at once.** Chernoff bounds S_t at one fixed t; UCB needed all $t \le T$ and paid a union bound. What is the right price for every t simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \dots, y_{t-1}$: the summands $x_t \eta_t$ are not independent and the fixed-design calculation is wrong.

Definition (where we stopped last time)

An adapted, integrable process M_t is a **martingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] = M_{t-1}$, and a **supermartingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] \le M_{t-1}$: given everything known after round $t-1$, on average it does not go up.

Two problems, one tool

1. **For all t at once.** Chernoff bounds S_t at one fixed t; UCB needed all $t \le T$ and paid a union bound. What is the right price for every t simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \dots, y_{t-1}$: the summands $x_t \eta_t$ are not independent and the fixed-design calculation is wrong.

Definition (where we stopped last time)

An adapted, integrable process M_t is a **martingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] = M_{t-1}$, and a **supermartingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] \le M_{t-1}$: given everything known after round $t-1$, on average it does not go up.

- $S_t = \sum_{s \le t} Z_s$ with $\mathbb{E}[Z_t \mid \mathcal{F}_{t-1}] = 0$ is a martingale: a fair game's running winnings. A real casino's is a supermartingale.
- $e^{\lambda S_t}$ is *not* one (Jensen pushes it up); the next slide corrects it into a supermartingale.

One supermartingale and one inequality will solve both problems: adaptivity by conditioning on $\mathcal{F}_{t-1}$, "all t" by never fixing t.

The supermartingale behind the Chernoff method

Let Z_t be adapted and conditionally 1-sub-Gaussian, $\mathbb{E}[e^{\lambda Z_t} \mid \mathcal{F}_{t-1}] \le e^{\lambda^2/2}$ for all $\lambda \in \mathbb{R}$, and $S_t = \sum_{s \le t} Z_s$. Fix λ and set

$$M_t(\lambda) := \exp\big(\lambda S_t - t\lambda^2/2\big), \qquad M_0(\lambda) = 1.$$

The supermartingale behind the Chernoff method

Let Z_t be adapted and conditionally 1-sub-Gaussian, $\mathbb{E}[e^{\lambda Z_t} \mid \mathcal{F}_{t-1}] \le e^{\lambda^2/2}$ for all $\lambda \in \mathbb{R}$, and $S_t = \sum_{s \le t} Z_s$. Fix λ and set

$$M_t(\lambda) := \exp\left(\lambda S_t - t\lambda^2/2\right), \qquad M_0(\lambda) = 1.$$

Peel off the last increment:

$$\mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] =$$

The supermartingale behind the Chernoff method

Let Z_t be adapted and conditionally 1-sub-Gaussian, $\mathbb{E}[e^{\lambda Z_t} \mid \mathcal{F}_{t-1}] \le e^{\lambda^2/2}$ for all $\lambda \in \mathbb{R}$, and $S_t = \sum_{s \le t} Z_s$. Fix λ and set

$$M_t(\lambda) := \exp\big(\lambda S_t - t\lambda^2/2\big), \qquad M_0(\lambda) = 1.$$

Peel off the last increment:

$$\mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] = M_{t-1}(\lambda)\, \mathbb{E}\big[e^{\lambda Z_t - \lambda^2/2} \mid \mathcal{F}_{t-1}\big] \;\le\; M_{t-1}(\lambda).$$

So $M_t(\lambda)$ is a nonnegative supermartingale starting at 1. Nothing about independence was used, only the conditional sub-Gaussian property, which is exactly what a bandit's noise satisfies given the past.

The supermartingale behind the Chernoff method

Let Z_t be adapted and conditionally 1-sub-Gaussian, $\mathbb{E}[e^{\lambda Z_t} \mid \mathcal{F}_{t-1}] \le e^{\lambda^2/2}$ for all $\lambda \in \mathbb{R}$, and $S_t = \sum_{s \le t} Z_s$. Fix λ and set

$$M_t(\lambda) := \exp\big(\lambda S_t - t\lambda^2/2\big), \qquad M_0(\lambda) = 1.$$

Peel off the last increment:

$$\mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] = M_{t-1}(\lambda)\, \mathbb{E}\big[e^{\lambda Z_t - \lambda^2/2} \mid \mathcal{F}_{t-1}\big] \;\le\; M_{t-1}(\lambda).$$

So $M_t(\lambda)$ is a nonnegative supermartingale starting at 1. Nothing about independence was used, only the conditional sub-Gaussian property, which is exactly what a bandit's noise satisfies given the past.

Where you have seen this

The Chernoff method bounds $\mathbb{P}(S_t \ge u)$ by $e^{-\lambda u}\mathbb{E}[e^{\lambda S_t}] \le e^{-\lambda u + t\lambda^2/2}$, i.e. by Markov applied to $M_t(\lambda)$ at a single time t. Tracking $M_t(\lambda)$ over time, and replacing Markov by its for-all-time version, is the whole idea of what follows.

Ville's inequality: Markov, for all time

Theorem (Ville 1939)

Let $M_t \geq 0$ be a supermartingale with $M_0 \leq 1$. Then for every $\delta \in (0, 1)$: $\mathbb{P}\big(\exists t \geq 0 : M_t \geq 1/\delta\big) \leq \delta$.

Markov at one fixed t gives $\mathbb{P}(M_t \geq 1/\delta) \leq \delta$. Ville says the same for the *whole path*: no union bound over t, no loss. Why, in three lines (proof: notes §1.3):

Ville's inequality: Markov, for all time

Theorem (Ville 1939)

Let $M_t \geq 0$ be a supermartingale with $M_0 \leq 1$. Then for every $\delta \in (0, 1)$: $\mathbb{P}\big(\exists t \geq 0 : M_t \geq 1/\delta\big) \leq \delta$.

Markov at one fixed t gives $\mathbb{P}(M_t \geq 1/\delta) \leq \delta$. Ville says the same for the *whole path*: no union bound over t, no loss. Why, in three lines (proof: notes §1.3):

- M_t is your fortune in a fair-or-unfavorable game, starting from \$1. "Do I *ever* reach $\$1/\delta$?" is about one stopping rule: quit the first time you do.

Ville's inequality: Markov, for all time

Theorem (Ville 1939)

Let $M_t \geq 0$ be a supermartingale with $M_0 \leq 1$. Then for every $\delta \in (0, 1)$: $\mathbb{P}\big(\exists t \geq 0 : M_t \geq 1/\delta\big) \leq \delta$.

Markov at one fixed t gives $\mathbb{P}(M_t \geq 1/\delta) \leq \delta$. Ville says the same for the *whole path*: no union bound over t, no loss. Why, in three lines (proof: notes §1.3):

- M_t is your fortune in a fair-or-unfavorable game, starting from \$1. "Do I *ever* reach $\$1/\delta$?" is about one stopping rule: quit the first time you do.
- Quitting based only on the past cannot make an unfavorable game favorable, so your expected fortune at quitting is still at most \$1.

Ville's inequality: Markov, for all time

Theorem (Ville 1939)

Let $M_t \geq 0$ be a supermartingale with $M_0 \leq 1$. Then for every $\delta \in (0, 1)$: $\mathbb{P}\big(\exists t \geq 0 : M_t \geq 1/\delta\big) \leq \delta$.

Markov at one fixed t gives $\mathbb{P}(M_t \geq 1/\delta) \leq \delta$. Ville says the same for the *whole path*: no union bound over t, no loss. Why, in three lines (proof: notes §1.3):

- M_t is your fortune in a fair-or-unfavorable game, starting from \$1. "Do I *ever* reach $\$1/\delta$?" is about one stopping rule: quit the first time you do.
- Quitting based only on the past cannot make an unfavorable game favorable, so your expected fortune at quitting is still at most \$1.
- If you ever reach $\$1/\delta$, you quit with at least $\$1/\delta$. Markov on the fortune at quitting time: probability at most δ.

Ville's inequality: Markov, for all time

Theorem (Ville 1939)

Let $M_t \geq 0$ be a supermartingale with $M_0 \leq 1$. Then for every $\delta \in (0, 1)$: $\mathbb{P}\big(\exists t \geq 0 : M_t \geq 1/\delta\big) \leq \delta$.

Markov at one fixed t gives $\mathbb{P}(M_t \geq 1/\delta) \leq \delta$. Ville says the same for the *whole path*: no union bound over t, no loss. Why, in three lines (proof: notes §1.3):

- M_t is your fortune in a fair-or-unfavorable game, starting from \$1. "Do I *ever* reach $\$1/\delta$?" is about one stopping rule: quit the first time you do.
- Quitting based only on the past cannot make an unfavorable game favorable, so your expected fortune at quitting is still at most \$1.
- If you ever reach $\$1/\delta$, you quit with at least $\$1/\delta$. Markov on the fortune at quitting time: probability at most δ.

How to use it

Build a nonnegative supermartingale with $M_0 = 1$ from the quantity you want to control and solve $M_t \geq 1/\delta$ for it: a boundary it stays under for all t at once, w.p. $1 - \delta$.

Linear boundaries

Apply Ville to $M_t(\lambda) = e^{\lambda S_t - t\lambda^2/2}$ for a fixed $\lambda > 0$:
$M_t(\lambda) \geq 1/\delta \iff S_t \geq$

Linear boundaries

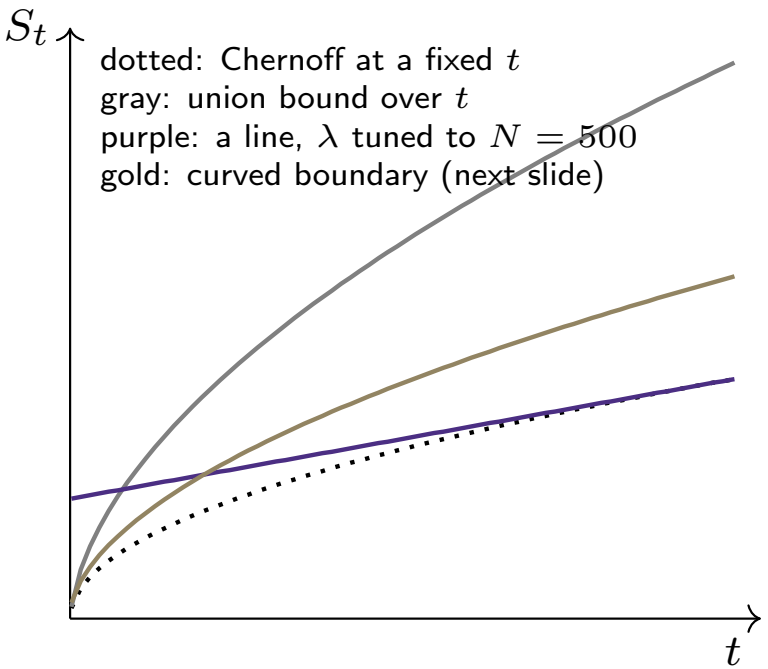
Apply Ville to $M_t(\lambda) = e^{\lambda S_t - t\lambda^2/2}$ for a fixed $\lambda > 0$:
$M_t(\lambda) \geq 1/\delta \iff S_t \geq \frac{t\lambda}{2} + \frac{\log(1/\delta)}{\lambda}$,

$$\text{so} \qquad \mathbb{P}\left(\exists t : \; S_t \geq \frac{t\lambda}{2} + \frac{\log(1/\delta)}{\lambda}\right) \leq \delta \qquad \text{for every fixed } \lambda > 0.$$

Linear boundaries

Apply Ville to $M_t(\lambda) = e^{\lambda S_t - t\lambda^2/2}$ for a fixed $\lambda > 0$:
$M_t(\lambda) \geq 1/\delta \iff S_t \geq \frac{t\lambda}{2} + \frac{\log(1/\delta)}{\lambda}$,

$$\text{so} \qquad \mathbb{P}\left(\exists t:\ S_t \geq \frac{t\lambda}{2} + \frac{\log(1/\delta)}{\lambda}\right) \leq \delta \qquad \text{for every fixed } \lambda > 0.$$



$\delta = 0.05$

- A line touches the $\sqrt{t}$ Chernoff curve at one time. With $\lambda = \sqrt{2\log(1/\delta)/N}$ it equals $\sqrt{2N\log(1/\delta)}$ at $t = N$: the fixed-time bound, now valid for *all* t, free at that one time.
- Away from N it is loose: linear for $t \gg N$, constant as $t \to 0$.
- Better than the union bound near N, but λ had to be chosen in advance.

Curved boundaries: the method of mixtures

Each $M_t(\lambda)$ is a supermartingale, so is any average of them: for a density h on $\mathbb{R}$, $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ has $\mathbb{E}[\bar{M}_t \mid \mathcal{F}_{t-1}] = \int \mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] h(\lambda) d\lambda \le \bar{M}_{t-1}$ and $\bar{M}_0 = 1$. One bet spread over many lines; Ville charges once.

Curved boundaries: the method of mixtures

Each $M_t(\lambda)$ is a supermartingale, so is any average of them: for a density h on $\mathbb{R}$, $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ has $\mathbb{E}[\bar{M}_t \mid \mathcal{F}_{t-1}] = \int \mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] h(\lambda) d\lambda \le \bar{M}_{t-1}$ and $\bar{M}_0 = 1$. One bet spread over many lines; Ville charges once. **Gaussian mixing**, $h(\lambda) = (2\pi\nu^2)^{-1/2} e^{-\lambda^2/(2\nu^2)}$; completing the square in λ (a Gaussian integral with variance $(t + \nu^{-2})^{-1}$),

$$\bar{M}_t = \frac{1}{\sqrt{2\pi\nu^2}} \int \exp\left(\lambda S_t - \frac{\lambda^2}{2}\left(t + \nu^{-2}\right)\right) d\lambda =$$

Curved boundaries: the method of mixtures

Each $M_t(\lambda)$ is a supermartingale, so is any average of them: for a density h on $\mathbb{R}$, $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ has $\mathbb{E}[\bar{M}_t \mid \mathcal{F}_{t-1}] = \int \mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] h(\lambda) d\lambda \le \bar{M}_{t-1}$ and $\bar{M}_0 = 1$. One bet spread over many lines; Ville charges once. **Gaussian mixing**, $h(\lambda) = (2\pi\nu^2)^{-1/2} e^{-\lambda^2/(2\nu^2)}$; completing the square in λ (a Gaussian integral with variance $(t + \nu^{-2})^{-1}$),

$$\bar{M}_t = \frac{1}{\sqrt{2\pi\nu^2}} \int \exp\Big(\lambda S_t - \frac{\lambda^2}{2}\big(t + \nu^{-2}\big)\Big) d\lambda = \sqrt{\frac{\nu^{-2}}{t + \nu^{-2}}}\, \exp\Big(\frac{S_t^2}{2(t + \nu^{-2})}\Big)$$

Curved boundaries: the method of mixtures

Each $M_t(\lambda)$ is a supermartingale, so is any average of them: for a density h on $\mathbb{R}$, $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ has $\mathbb{E}[\bar{M}_t \mid \mathcal{F}_{t-1}] = \int \mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] h(\lambda) d\lambda \le \bar{M}_{t-1}$ and $\bar{M}_0 = 1$. One bet spread over many lines; Ville charges once. **Gaussian mixing**, $h(\lambda) = (2\pi\nu^2)^{-1/2} e^{-\lambda^2/(2\nu^2)}$; completing the square in λ (a Gaussian integral with variance $(t+\nu^{-2})^{-1}$),

$$\bar{M}_t = \frac{1}{\sqrt{2\pi\nu^2}} \int \exp\Big(\lambda S_t - \frac{\lambda^2}{2}\big(t+\nu^{-2}\big)\Big) d\lambda = \sqrt{\frac{\nu^{-2}}{t+\nu^{-2}}}\, \exp\Big(\frac{S_t^2}{2(t+\nu^{-2})}\Big)$$

Ville, and solve $\bar{M}_t \ge 1/\delta$ for $|S_t|$:

$$\mathbb{P}\Big(\exists t : \ |S_t| \ge \qquad\qquad\qquad\qquad\qquad\qquad \Big) \le \delta.$$

Curved boundaries: the method of mixtures

Each $M_t(\lambda)$ is a supermartingale, so is any average of them: for a density h on $\mathbb{R}$, $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ has $\mathbb{E}[\bar{M}_t \mid \mathcal{F}_{t-1}] = \int \mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] h(\lambda) d\lambda \le \bar{M}_{t-1}$ and $\bar{M}_0 = 1$. One bet spread over many lines; Ville charges once. **Gaussian mixing**, $h(\lambda) = (2\pi\nu^2)^{-1/2} e^{-\lambda^2/(2\nu^2)}$; completing the square in λ (a Gaussian integral with variance $(t+\nu^{-2})^{-1}$),

$$\bar{M}_t = \frac{1}{\sqrt{2\pi\nu^2}} \int \exp\Big(\lambda S_t - \frac{\lambda^2}{2}\big(t+\nu^{-2}\big)\Big) d\lambda = \sqrt{\frac{\nu^{-2}}{t+\nu^{-2}}}\, \exp\Big(\frac{S_t^2}{2(t+\nu^{-2})}\Big)$$

Ville, and solve $\bar{M}_t \ge 1/\delta$ for $|S_t|$:

$$\mathbb{P}\Big(\exists t:\ |S_t| \ge \sqrt{2\big(t+\nu^{-2}\big)\Big(\log\tfrac{1}{\delta} + \tfrac{1}{2}\log\big((t+\nu^{-2})\nu^2\big)\Big)}\Big) \le \delta.$$

With $\nu = 1$: $|S_t| \le \sqrt{2(t+1)\log(\sqrt{t+1}/\delta)}$ for all t. Compare the union bound $\sqrt{2t\log(2t^2/\delta)}$: the $\log t^2$ became $\frac{1}{2}\log t$, and the curve is tightest around $t \approx \nu^{-2}$, which is where h put its mass. (Law of the iterated logarithm: the truth is $\sqrt{2t\log\log t}$; a cleverer h gets it.)

What $M_t(\lambda)$ is: a likelihood ratio

Take Gaussian noise, $Z_s \sim \mathcal{N}(\mu, 1)$ i.i.d., and the binary test $H_0 : \mu = 0$ against $H_\lambda : \mu = \lambda$. The likelihood ratio after t samples is

$$L_t(\lambda) := \prod_{s=1}^{t} \frac{p_\lambda(Z_s)}{p_0(Z_s)} = \prod_{s=1}^{t} \frac{e^{-(Z_s-\lambda)^2/2}}{e^{-Z_s^2/2}} =$$

What $M_t(\lambda)$ is: a likelihood ratio

Take Gaussian noise, $Z_s \sim \mathcal{N}(\mu, 1)$ i.i.d., and the binary test $H_0 : \mu = 0$ against $H_\lambda : \mu = \lambda$. The likelihood ratio after t samples is

$$L_t(\lambda) := \prod_{s=1}^{t} \frac{p_\lambda(Z_s)}{p_0(Z_s)} = \prod_{s=1}^{t} \frac{e^{-(Z_s-\lambda)^2/2}}{e^{-Z_s^2/2}} = \prod_{s=1}^{t} e^{\lambda Z_s - \lambda^2/2} = e^{\lambda S_t - t\lambda^2/2} = M_t(\lambda).$$

So $M_t(\lambda)$ is the *evidence for the alternative λ against the null*, and under the null it is a martingale for *any* pair of densities: $\mathbb{E}_0\big[p_\lambda(Z)/p_0(Z)\big] = \int p_\lambda = 1$.

What $M_t(\lambda)$ is: a likelihood ratio

Take Gaussian noise, $Z_s \sim \mathcal{N}(\mu, 1)$ i.i.d., and the binary test $H_0 : \mu = 0$ against $H_\lambda : \mu = \lambda$. The likelihood ratio after t samples is

$$L_t(\lambda) := \prod_{s=1}^{t} \frac{p_\lambda(Z_s)}{p_0(Z_s)} = \prod_{s=1}^{t} \frac{e^{-(Z_s-\lambda)^2/2}}{e^{-Z_s^2/2}} = \prod_{s=1}^{t} e^{\lambda Z_s - \lambda^2/2} = e^{\lambda S_t - t\lambda^2/2} = M_t(\lambda).$$

So $M_t(\lambda)$ is the *evidence for the alternative λ against the null*, and under the null it is a martingale for *any* pair of densities: $\mathbb{E}_0\big[p_\lambda(Z)/p_0(Z)\big] = \int p_\lambda = 1$.

Ville, restated

Under the true hypothesis, the likelihood ratio in favor of any fixed alternative exceeds $1/\delta$ at *some* time with probability at most δ, however long you keep sampling.

This is the statement behind sequential testing: you may look at the data after every sample and stop whenever you like, and "the evidence against the truth is $1/\delta$-strong" still happens with probability $\leq \delta$. No union bound over stopping times is needed, because the likelihood ratio is the martingale.

For fun: the sequential probability ratio test

Densities p_0, p_1, i.i.d. samples, $L_t = \prod_{s \le t} p_1(Z_s)/p_0(Z_s)$.

SPRT (Wald 1945), one-sided

Sample until $\tau = \min\{t : L_t \ge 1/\delta\}$, then declare H_1. By Ville, $\mathbb{P}_0(\text{declare } H_1) = \mathbb{P}_0(\exists t : L_t \ge 1/\delta) \le \delta$.

For fun: the sequential probability ratio test

Densities p_0, p_1, i.i.d. samples, $L_t = \prod_{s \le t} p_1(Z_s)/p_0(Z_s)$.

SPRT (Wald 1945), one-sided

Sample until $\tau = \min\{t : L_t \ge 1/\delta\}$, then declare H_1. By Ville, $\mathbb{P}_0(\text{declare } H_1) = \mathbb{P}_0(\exists t : L_t \ge 1/\delta) \le \delta$.

Wald's identity (notes Lemma 4)

$W_1, W_2, \ldots$ i.i.d. with mean m, τ a stopping time with $\mathbb{E}[\tau] < \infty$: $\mathbb{E}\big[\sum_{s \le \tau} W_s\big] = m\,\mathbb{E}[\tau]$.

For fun: the sequential probability ratio test

Densities p_0, p_1, i.i.d. samples, $L_t = \prod_{s \le t} p_1(Z_s)/p_0(Z_s)$.

SPRT (Wald 1945), one-sided

Sample until $\tau = \min\{t : L_t \ge 1/\delta\}$, then declare H_1. By Ville, $\mathbb{P}_0(\text{declare } H_1) = \mathbb{P}_0(\exists t : L_t \ge 1/\delta) \le \delta$.

Wald's identity (notes Lemma 4)

$W_1, W_2, \ldots$ i.i.d. with mean m, τ a stopping time with $\mathbb{E}[\tau] < \infty$: $\mathbb{E}\big[\sum_{s \le \tau} W_s\big] = m\,\mathbb{E}[\tau]$.

Under H_1 take $W_s = \log \frac{p_1(Z_s)}{p_0(Z_s)}$, mean $\mathbb{E}_1[W_s] = \mathrm{KL}(p_1 \| p_0)$; assume $W_s \le \log \kappa$ (bounded overshoot):

$$\log \tfrac{1}{\delta} \;\le\; \log L_\tau = \sum_{s \le \tau} W_s \;\le\; \log \tfrac{1}{\delta} + \log \kappa,$$

so, taking $\mathbb{E}_1$: $\quad \log \tfrac{1}{\delta} \;\le\; \le\; \log \tfrac{\kappa}{\delta}.$

For fun: the sequential probability ratio test

Densities p_0, p_1, i.i.d. samples, $L_t = \prod_{s \le t} p_1(Z_s)/p_0(Z_s)$.

SPRT (Wald 1945), one-sided

Sample until $\tau = \min\{t : L_t \ge 1/\delta\}$, then declare H_1. By Ville, $\mathbb{P}_0(\text{declare } H_1) = \mathbb{P}_0(\exists t : L_t \ge 1/\delta) \le \delta$.

Wald's identity (notes Lemma 4)

$W_1, W_2, \ldots$ i.i.d. with mean m, τ a stopping time with $\mathbb{E}[\tau] < \infty$: $\mathbb{E}\big[\sum_{s \le \tau} W_s\big] = m\,\mathbb{E}[\tau]$.

Under H_1 take $W_s = \log \frac{p_1(Z_s)}{p_0(Z_s)}$, mean $\mathbb{E}_1[W_s] = \mathrm{KL}(p_1 \| p_0)$; assume $W_s \le \log \kappa$ (bounded overshoot):

$$\log \tfrac{1}{\delta} \;\le\; \log L_\tau = \sum_{s \le \tau} W_s \;\le\; \log \tfrac{1}{\delta} + \log \kappa,$$

so, taking $\mathbb{E}_1$:

$$\log \tfrac{1}{\delta} \;\le\; \mathrm{KL}(p_1 \| p_0)\, \mathbb{E}_1[\tau] \;\le\; \log \tfrac{\kappa}{\delta}.$$

So $\mathbb{E}_1[\tau] \approx \log(1/\delta)/\mathrm{KL}(p_1 \| p_0)$: the evidence needed, divided by the evidence per sample.

And the SPRT is optimal: Lecture 8's template

Any sequential test (stop at τ, declare a hypothesis) with both error probabilities at most δ:

And the SPRT is optimal: Lecture 8's template

Any sequential test (stop at τ, declare a hypothesis) with both error probabilities at most δ:

- **Bretagnolle–Huber** with $A = \{\text{declare } H_0\}$:
  $2\delta \;\geq\; \mathbb{P}_1(\text{declare } H_0) + \mathbb{P}_0(\text{declare } H_1) \;\geq\; \frac{1}{2}\exp\big(-\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau)\big)$, so $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) \geq \log\frac{1}{4\delta}$.

And the SPRT is optimal: Lecture 8's template

Any sequential test (stop at τ, declare a hypothesis) with both error probabilities at most δ:

- **Bretagnolle–Huber** with $A = \{\text{declare } H_0\}$:
  $2\delta \;\geq\; \mathbb{P}_1(\text{declare } H_0) + \mathbb{P}_0(\text{declare } H_1) \;\geq\; \frac{1}{2}\exp\big(-\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau)\big)$, so $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) \geq \log\frac{1}{4\delta}$.
- **Divergence decomposition at a stopping time:**
  $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) = \mathbb{E}_1\big[\sum_{s\leq\tau} \log\frac{p_1(Z_s)}{p_0(Z_s)}\big] = \mathbb{E}_1[\tau]\,\mathrm{KL}(p_1\|p_0)$, which is Wald's identity. (Fact 1 of Lecture 8 was this, with arms.)

And the SPRT is optimal: Lecture 8's template

Any sequential test (stop at τ, declare a hypothesis) with both error probabilities at most δ:

- **Bretagnolle–Huber** with $A = \{\text{declare } H_0\}$: $2\delta \;\geq\; \mathbb{P}_1(\text{declare } H_0) + \mathbb{P}_0(\text{declare } H_1) \;\geq\; \frac{1}{2}\exp\big(-\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau)\big)$, so $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) \geq \log\frac{1}{4\delta}$.
- **Divergence decomposition at a stopping time:** $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) = \mathbb{E}_1\big[\sum_{s\leq\tau} \log\frac{p_1(Z_s)}{p_0(Z_s)}\big] = \mathbb{E}_1[\tau]\,\mathrm{KL}(p_1\|p_0)$, which is Wald's identity. (Fact 1 of Lecture 8 was this, with arms.)
- **Hence** $\mathbb{E}_1[\tau] \geq \log(1/(4\delta))/\mathrm{KL}(p_1\|p_0)$, and the SPRT matches it up to the constant inside the log. For Gaussians with gap Δ, $\mathrm{KL} = \Delta^2/2$: about $2\log(1/\delta)/\Delta^2$ samples, Lecture 6's A/B rate with the right constant and no horizon.

And the SPRT is optimal: Lecture 8's template

Any sequential test (stop at τ, declare a hypothesis) with both error probabilities at most δ:

- **Bretagnolle–Huber** with $A = \{\text{declare } H_0\}$: $2\delta \;\geq\; \mathbb{P}_1(\text{declare } H_0) + \mathbb{P}_0(\text{declare } H_1) \;\geq\; \frac{1}{2}\exp\big(-\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau)\big)$, so $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) \geq \log\frac{1}{4\delta}$.
- **Divergence decomposition at a stopping time:** $\mathrm{KL}(\mathbb{P}_1^\tau, \mathbb{P}_0^\tau) = \mathbb{E}_1\big[\sum_{s\leq\tau} \log\frac{p_1(Z_s)}{p_0(Z_s)}\big] = \mathbb{E}_1[\tau]\,\mathrm{KL}(p_1\|p_0)$, which is Wald's identity. (Fact 1 of Lecture 8 was this, with arms.)
- **Hence** $\mathbb{E}_1[\tau] \geq \log(1/(4\delta))/\mathrm{KL}(p_1\|p_0)$, and the SPRT matches it up to the constant inside the log. For Gaussians with gap Δ, $\mathrm{KL} = \Delta^2/2$: about $2\log(1/\delta)/\Delta^2$ samples, Lecture 6's A/B rate with the right constant and no horizon.

The moral

One object, two jobs: tracked as a martingale and stopped, the likelihood ratio is an optimal test; its expectation under the alternative is the information any test must collect.

What $\bar{M}_t$ is: a Bayes factor, with a prior on the gap

A single alternative λ is a bet on one effect size. The mixture

$$\bar{M}_t = \int \prod_{s \le t} \frac{p_\lambda(Z_s)}{p_0(Z_s)}\, h(\lambda)\, d\lambda = \frac{\int p_\lambda(Z_1, \ldots, Z_t)\, h(\lambda)\, d\lambda}{p_0(Z_1, \ldots, Z_t)}$$

is the likelihood ratio of the *composite* alternative "the gap μ is drawn from h" against $\mu = 0$: a **Bayes factor**, with h a prior on the effect size.

What $\bar{M}_t$ is: a Bayes factor, with a prior on the gap

A single alternative λ is a bet on one effect size. The mixture

$$\bar{M}_t = \int \prod_{s \le t} \frac{p_\lambda(Z_s)}{p_0(Z_s)}\, h(\lambda)\, d\lambda = \frac{\int p_\lambda(Z_1, \ldots, Z_t)\, h(\lambda)\, d\lambda}{p_0(Z_1, \ldots, Z_t)}$$

is the likelihood ratio of the *composite* alternative "the gap μ is drawn from h" against $\mu = 0$: a **Bayes factor**, with h a prior on the effect size.

- It is still a martingale under the null (an average of martingales), so Ville applies: the Bayes factor against the truth rarely exceeds $1/\delta$.
- $h = \mathcal{N}(0, \nu^2)$ says "I expect a gap of order ν". The boundary is tightest near $t \approx \nu^{-2}$, where such a gap becomes detectable, and only mildly worse elsewhere: the prior hedges across effect sizes the way a union bound hedges across times, but smoothly.

What $\bar{M}_t$ is: a Bayes factor, with a prior on the gap

A single alternative λ is a bet on one effect size. The mixture

$$\bar{M}_t = \int \prod_{s \le t} \frac{p_\lambda(Z_s)}{p_0(Z_s)}\, h(\lambda)\, d\lambda = \frac{\int p_\lambda(Z_1, \ldots, Z_t)\, h(\lambda)\, d\lambda}{p_0(Z_1, \ldots, Z_t)}$$

is the likelihood ratio of the *composite* alternative "the gap μ is drawn from h" against $\mu = 0$: a **Bayes factor**, with h a prior on the effect size.

- It is still a martingale under the null (an average of martingales), so Ville applies: the Bayes factor against the truth rarely exceeds $1/\delta$.
- $h = \mathcal{N}(0, \nu^2)$ says "I expect a gap of order ν". The boundary is tightest near $t \approx \nu^{-2}$, where such a gap becomes detectable, and only mildly worse elsewhere: the prior hedges across effect sizes the way a union bound hedges across times, but smoothly.
- **In $\mathbb{R}^d$:** with $y_s = \langle x_s, \theta \rangle + \eta_s$ the likelihood ratio of θ against 0 is $\exp(\langle \theta, S_t \rangle - \frac{1}{2} \|\theta\|^2_{V_t}) = M_t(\theta)$; mixing with the prior $\theta \sim \mathcal{N}(0, \gamma^{-1} I)$ gives the Bayes factor we compute next. Ridge regression is the posterior mean under this prior and γ its precision.

Vector-valued martingales: the bandit's noise process

Now $x_t \in \mathbb{R}^d$ predictable (the arm chosen from the past), η_t conditionally 1-sub-Gaussian, and $Z_t = x_t \eta_t$, $S_t = \sum_{s \le t} x_s \eta_s$, $V_t = \sum_{s \le t} x_s x_s^\top$. For a fixed $\lambda \in \mathbb{R}^d$, $\mathbb{E}\left[e^{\langle \lambda, Z_t \rangle} \mid \mathcal{F}_{t-1}\right] = \mathbb{E}\left[e^{\langle \lambda, x_t \rangle \eta_t} \mid \mathcal{F}_{t-1}\right] \le$

Vector-valued martingales: the bandit's noise process

Now $x_t \in \mathbb{R}^d$ predictable (the arm chosen from the past), η_t conditionally 1-sub-Gaussian, and $Z_t = x_t \eta_t$, $S_t = \sum_{s \le t} x_s \eta_s$, $V_t = \sum_{s \le t} x_s x_s^\top$. For a fixed $\lambda \in \mathbb{R}^d$, $\mathbb{E}\big[e^{\langle \lambda, Z_t \rangle} \mid \mathcal{F}_{t-1}\big] = \mathbb{E}\big[e^{\langle \lambda, x_t \rangle \eta_t} \mid \mathcal{F}_{t-1}\big] \le e^{\langle \lambda, x_t \rangle^2 / 2}$, because $\langle \lambda, x_t \rangle$ is a constant given $\mathcal{F}_{t-1}$, and $\langle \lambda, x_t \rangle^2 = \|\lambda\|^2_{x_t x_t^\top}$ adds up to $\|\lambda\|^2_{V_t}$. Exactly as before,

$$M_t(\lambda) = \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \right)$$

is a nonnegative supermartingale with $M_0(\lambda) = 1$, and Ville gives a *linear* boundary for each fixed λ: $\mathbb{P}\big(\exists t : \langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \ge \log \tfrac{1}{\delta}\big) \le \delta$.

Vector-valued martingales: the bandit's noise process

Now $x_t \in \mathbb{R}^d$ predictable (the arm chosen from the past), η_t conditionally 1-sub-Gaussian, and $Z_t = x_t \eta_t$, $S_t = \sum_{s \le t} x_s \eta_s$, $V_t = \sum_{s \le t} x_s x_s^\top$. For a fixed $\lambda \in \mathbb{R}^d$, $\mathbb{E}\big[e^{\langle \lambda, Z_t \rangle} \mid \mathcal{F}_{t-1}\big] = \mathbb{E}\big[e^{\langle \lambda, x_t \rangle \eta_t} \mid \mathcal{F}_{t-1}\big] \le e^{\langle \lambda, x_t \rangle^2 / 2}$, because $\langle \lambda, x_t \rangle$ is a constant given $\mathcal{F}_{t-1}$, and $\langle \lambda, x_t \rangle^2 = \|\lambda\|^2_{x_t x_t^\top}$ adds up to $\|\lambda\|^2_{V_t}$. Exactly as before,

$$M_t(\lambda) = \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \right)$$

is a nonnegative supermartingale with $M_0(\lambda) = 1$, and Ville gives a *linear* boundary for each fixed λ: $\mathbb{P}\big(\exists t : \langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \ge \log \tfrac{1}{\delta}\big) \le \delta$.

The same obstacle as in the fixed design

We want $\|S_t\|_{V_t^{-1}}$, and $\max_\lambda \left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \right) = \tfrac{1}{2} \|S_t\|^2_{V_t^{-1}}$ at $\lambda = V_t^{-1} S_t$. But that λ is random, and the boundary holds for fixed λ only. Last time: a net. Here: a better move.

Vector-valued martingales: the bandit's noise process

Now $x_t \in \mathbb{R}^d$ predictable (the arm chosen from the past), η_t conditionally 1-sub-Gaussian, and $Z_t = x_t \eta_t$, $S_t = \sum_{s \le t} x_s \eta_s$, $V_t = \sum_{s \le t} x_s x_s^\top$. For a fixed $\lambda \in \mathbb{R}^d$, $\mathbb{E}\big[e^{\langle \lambda, Z_t \rangle} \mid \mathcal{F}_{t-1}\big] = \mathbb{E}\big[e^{\langle \lambda, x_t \rangle \eta_t} \mid \mathcal{F}_{t-1}\big] \le e^{\langle \lambda, x_t \rangle^2 / 2}$, because $\langle \lambda, x_t \rangle$ is a constant given $\mathcal{F}_{t-1}$, and $\langle \lambda, x_t \rangle^2 = \|\lambda\|^2_{x_t x_t^\top}$ adds up to $\|\lambda\|^2_{V_t}$. Exactly as before,

$$M_t(\lambda) = \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \right)$$

is a nonnegative supermartingale with $M_0(\lambda) = 1$, and Ville gives a *linear* boundary for each fixed λ: $\mathbb{P}\big(\exists t : \langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \ge \log \tfrac{1}{\delta}\big) \le \delta$.

The same obstacle as in the fixed design

We want $\|S_t\|_{V_t^{-1}}$, and $\max_\lambda \left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|^2_{V_t} \right) = \tfrac{1}{2} \|S_t\|^2_{V_t^{-1}}$ at $\lambda = V_t^{-1} S_t$. But that λ is random, and the boundary holds for fixed λ only. Last time: a net. Here: a better move.

Mix over λ: $h = \mathcal{N}(0, \gamma^{-1} I_d)$. Then $\bar{M}_t = \int M_t(\lambda) h(\lambda) d\lambda$ is a supermartingale with $\bar{M}_0 = 1$, and the integral is Gaussian.

The mixture, computed

With $h(\lambda) = (\gamma/2\pi)^{d/2} e^{-\gamma \|\lambda\|_2^2/2}$, write $\bar{V}_t = V_t + \gamma I$ and complete the square:

$$\begin{aligned} \bar{M}_t &= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2} \|\lambda\|_{\bar{V}_t}^2\right) d\lambda \\ &= \\ &= \end{aligned}$$

The mixture, computed

With $h(\lambda) = (\gamma/2\pi)^{d/2} e^{-\gamma\|\lambda\|_2^2/2}$, write $\bar{V}_t = V_t + \gamma I$ and complete the square:

$$\begin{aligned}
\bar{M}_t &= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2}\|\lambda\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2 - \tfrac{1}{2}\left\|\lambda - \bar{V}_t^{-1} S_t\right\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} (2\pi)^{d/2} \det(\bar{V}_t)^{-1/2} \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2\right) = \left(\frac{\gamma^d}{\det \bar{V}_t}\right)^{1/2} e^{\frac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2}.
\end{aligned}$$

(The λ-integral is Gaussian with covariance $\bar{V}_t^{-1}$.) Now Ville: $\bar{M}_t \geq 1/\delta$ iff

$$\|S_t\|_{\bar{V}_t^{-1}}^2 \;\geq$$

The mixture, computed

With $h(\lambda) = (\gamma/2\pi)^{d/2} e^{-\gamma\|\lambda\|_2^2/2}$, write $\bar{V}_t = V_t + \gamma I$ and complete the square:

$$\begin{aligned}
\bar{M}_t &= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\langle \lambda, S_t \rangle - \tfrac{1}{2}\|\lambda\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2 - \tfrac{1}{2}\left\|\lambda - \bar{V}_t^{-1} S_t\right\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} (2\pi)^{d/2} \det(\bar{V}_t)^{-1/2} \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2\right) = \left(\frac{\gamma^d}{\det \bar{V}_t}\right)^{1/2} e^{\frac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2}.
\end{aligned}$$

(The λ-integral is Gaussian with covariance $\bar{V}_t^{-1}$.) Now Ville: $\bar{M}_t \geq 1/\delta$ iff

$$\|S_t\|_{\bar{V}_t^{-1}}^2 \;\geq\; 2\log\frac{1}{\delta} + \log\frac{\det \bar{V}_t}{\gamma^d}.$$

The mixture, computed

With $h(\lambda) = (\gamma/2\pi)^{d/2} e^{-\gamma\|\lambda\|_2^2/2}$, write $\bar{V}_t = V_t + \gamma I$ and complete the square:

$$\begin{aligned}
\bar{M}_t &= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\langle \lambda, S_t\rangle - \tfrac{1}{2}\|\lambda\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} \int \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2 - \tfrac{1}{2}\left\|\lambda - \bar{V}_t^{-1} S_t\right\|_{\bar{V}_t}^2\right) d\lambda \\
&= \left(\frac{\gamma}{2\pi}\right)^{d/2} (2\pi)^{d/2} \det(\bar{V}_t)^{-1/2} \exp\left(\tfrac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2\right) = \left(\frac{\gamma^d}{\det \bar{V}_t}\right)^{1/2} e^{\frac{1}{2}\|S_t\|_{\bar{V}_t^{-1}}^2}.
\end{aligned}$$

(The λ-integral is Gaussian with covariance $\bar{V}_t^{-1}$.) Now Ville: $\bar{M}_t \geq 1/\delta$ iff

$$\|S_t\|_{\bar{V}_t^{-1}}^2 \;\geq\; 2\log\frac{1}{\delta} + \log\frac{\det \bar{V}_t}{\gamma^d}.$$

Self-normalized bound (Abbasi-Yadkori et al. 2011; notes eq. 1.1)

$$\mathbb{P}\left(\exists t :\; \|S_t\|_{(V_t+\gamma I)^{-1}} \geq \sqrt{2\log\tfrac{1}{\delta} + \log\tfrac{\det(V_t+\gamma I)}{\gamma^d}}\right) \leq \delta.$$

Sanity check: $d = 1$, $x_t \equiv 1$, $\gamma = \nu^{-2}$ gives $V_t = t$ and exactly the curved boundary from before.

Finale: the confidence ellipsoid, proved

Ridge with parameter γ (called λ last lecture): $\hat{\theta}_t = \bar{V}_t^{-1} \sum_{s \le t} x_s y_s$, $\bar{V}_t = V_t + \gamma I$. Substitute $y_s = \langle x_s, \theta^* \rangle + \eta_s$:

$$\hat{\theta}_t = \bar{V}_t^{-1} \big(V_t \theta^* + S_t \big) =$$

so $$\left\| \hat{\theta}_t - \theta^* \right\|_{\bar{V}_t} =$$

Finale: the confidence ellipsoid, proved

Ridge with parameter γ (called λ last lecture): $\hat{\theta}_t = \bar{V}_t^{-1} \sum_{s \le t} x_s y_s$, $\bar{V}_t = V_t + \gamma I$. Substitute $y_s = \langle x_s, \theta^* \rangle + \eta_s$:

$$\hat{\theta}_t = \bar{V}_t^{-1}\big(V_t\theta^* + S_t\big) = \theta^* + \bar{V}_t^{-1}\big(S_t - \gamma\theta^*\big),$$

so $$\left\|\hat{\theta}_t - \theta^*\right\|_{\bar{V}_t} = \|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}}\,.$$

Finale: the confidence ellipsoid, proved

Ridge with parameter γ (called λ last lecture): $\hat{\theta}_t = \bar{V}_t^{-1} \sum_{s \le t} x_s y_s$, $\bar{V}_t = V_t + \gamma I$. Substitute $y_s = \langle x_s, \theta^* \rangle + \eta_s$:

$$\hat{\theta}_t = \bar{V}_t^{-1}\big(V_t \theta^* + S_t\big) = \theta^* + \bar{V}_t^{-1}\big(S_t - \gamma\theta^*\big),$$

$$\text{so} \qquad \left\|\hat{\theta}_t - \theta^*\right\|_{\bar{V}_t} = \|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}}\,.$$

Triangle inequality, then $\bar{V}_t \succeq \gamma I$:
$\|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}} \le$

Finale: the confidence ellipsoid, proved

Ridge with parameter γ (called λ last lecture): $\hat{\theta}_t = \bar{V}_t^{-1} \sum_{s \le t} x_s y_s$, $\bar{V}_t = V_t + \gamma I$. Substitute $y_s = \langle x_s, \theta^* \rangle + \eta_s$:

$$\hat{\theta}_t = \bar{V}_t^{-1}\big(V_t \theta^* + S_t\big) = \theta^* + \bar{V}_t^{-1}\big(S_t - \gamma\theta^*\big),$$

$$\text{so} \qquad \left\|\hat{\theta}_t - \theta^*\right\|_{\bar{V}_t} = \|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}}\,.$$

Triangle inequality, then $\bar{V}_t \succeq \gamma I$:

$$\|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}} \le \|S_t\|_{\bar{V}_t^{-1}} + \gamma\,\|\theta^*\|_{\bar{V}_t^{-1}} \le \|S_t\|_{\bar{V}_t^{-1}} + \sqrt{\gamma}\,\|\theta^*\|_2.$$

Finale: the confidence ellipsoid, proved

Ridge with parameter γ (called λ last lecture): $\hat{\theta}_t = \bar{V}_t^{-1} \sum_{s \le t} x_s y_s$, $\bar{V}_t = V_t + \gamma I$. Substitute $y_s = \langle x_s, \theta^* \rangle + \eta_s$:

$$\hat{\theta}_t = \bar{V}_t^{-1}\big(V_t\theta^* + S_t\big) = \theta^* + \bar{V}_t^{-1}\big(S_t - \gamma\theta^*\big),$$

$$\text{so} \qquad \left\|\hat{\theta}_t - \theta^*\right\|_{\bar{V}_t} = \|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}}\,.$$

Triangle inequality, then $\bar{V}_t \succeq \gamma I$:
$\|S_t - \gamma\theta^*\|_{\bar{V}_t^{-1}} \le \|S_t\|_{\bar{V}_t^{-1}} + \gamma\,\|\theta^*\|_{\bar{V}_t^{-1}} \le \|S_t\|_{\bar{V}_t^{-1}} + \sqrt{\gamma}\;\|\theta^*\|_2$.

Theorem (the confidence ellipsoid, proved)

If $\|\theta^*\|_2 \le S$, then with probability at least $1-\delta$, simultaneously for all $t \ge 0$,

$$\begin{aligned}\left\|\hat{\theta}_t - \theta^*\right\|_{V_t+\gamma I} &\le \sqrt{\gamma}\,S + \sqrt{2\log\tfrac{1}{\delta} + \log\tfrac{\det(V_t+\gamma I)}{\gamma^d}} \\ &\le \sqrt{\gamma}\,S + \sqrt{2\log\tfrac{1}{\delta} + d\log\big(1 + \tfrac{tL^2}{d\gamma}\big)}.\end{aligned}$$

The last step: $\log\det\bar{V}_t \le d\log(\mathrm{tr}\bar{V}_t/d)$ (AM–GM on the eigenvalues) and $\mathrm{tr}\bar{V}_t \le d\gamma + tL^2$ when $\|x_s\|_2 \le L$. ■

Back to LinUCB: this is why we needed vector martingales

LinUCB plays $x_t = \arg\max_x \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$, using *all* past data, with x_t chosen *because of* past noise. Its analysis needed exactly one thing we could not prove at the time:

The good event

$\theta^* \in \mathcal{C}_t = \{\theta : \left\|\theta - \hat{\theta}_t\right\|_{V_t + \gamma I} \leq \beta_t\}$ for every t simultaneously, with probability $1 - \delta$.

Back to LinUCB: this is why we needed vector martingales

LinUCB plays $x_t = \arg\max_x \langle x, \hat{\theta}_{t-1} \rangle + \beta_{t-1} \|x\|_{V_{t-1}^{-1}}$, using *all* past data, with x_t chosen *because of* past noise. Its analysis needed exactly one thing we could not prove at the time:

The good event

$\theta^* \in \mathcal{C}_t = \{\theta : \left\|\theta - \hat{\theta}_t\right\|_{V_t + \gamma I} \leq \beta_t\}$ for every t simultaneously, with probability $1 - \delta$.

- **Adaptive**, because x_t depends on $y_1, \ldots, y_{t-1}$, so the fixed-design ellipsoid does not apply. The supermartingale $M_t(\lambda)$ only needed $\langle \lambda, x_t \rangle$ known before η_t arrives: exactly what a bandit algorithm guarantees.
- **For all t at once**, because the algorithm consults $\mathcal{C}_{t-1}$ every round and each consultation must be right: Ville, not a union bound over T rounds.
- **Every direction at once**, because the arm set may be infinite or change: the mixture over λ, not a union bound over arms; the $\sqrt{d \log t}$ radius is its price.

The LinUCB proof, now complete

1. **The good event** $\theta^* \in \mathcal{C}_t$ for all t holds with probability $1 - \delta$: today's theorem, via the vector supermartingale, Ville, and the Gaussian mixture.

The LinUCB proof, now complete

1. **The good event** $\theta^* \in \mathcal{C}_t$ for all t holds with probability $1 - \delta$: today's theorem, via the vector supermartingale, Ville, and the Gaussian mixture.
2. **One round**, on the good event: optimism and Cauchy–Schwarz, exactly as in the multi-armed UCB proof,

$$r_t \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \;\le\; 2\beta_{t-1} \, \|x_t\|_{V_{t-1}^{-1}} \, .$$

The LinUCB proof, now complete

1. **The good event** $\theta^* \in \mathcal{C}_t$ for all t holds with probability $1-\delta$: today's theorem, via the vector supermartingale, Ville, and the Gaussian mixture.
2. **One round**, on the good event: optimism and Cauchy–Schwarz, exactly as in the multi-armed UCB proof,

$$r_t \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \;\le\; 2\beta_{t-1}\,\|x_t\|_{V_{t-1}^{-1}}\,.$$

3. **Summing the widths**: the elliptical potential lemma, $\sum_t \min\{1, \|x_t\|^2_{V_{t-1}^{-1}}\} \le 2d\log(1 + TL^2/(d\gamma))$, which is linear algebra about any arm sequence.

The LinUCB proof, now complete

1. **The good event** $\theta^* \in \mathcal{C}_t$ for all t holds with probability $1-\delta$: today's theorem, via the vector supermartingale, Ville, and the Gaussian mixture.
2. **One round**, on the good event: optimism and Cauchy–Schwarz, exactly as in the multi-armed UCB proof,

$$r_t \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \;\le\; 2\beta_{t-1}\, \|x_t\|_{V_{t-1}^{-1}} .$$

3. **Summing the widths**: the elliptical potential lemma, $\sum_t \min\{1, \|x_t\|^2_{V_{t-1}^{-1}}\} \le 2d \log(1 + TL^2/(d\gamma))$, which is linear algebra about any arm sequence.
4. **Cauchy–Schwarz** glues them, so with probability $1-\delta$,

$$R_T \;\le\; 2\beta_T \sqrt{2dT \log\left(1 + \tfrac{TL^2}{d\gamma}\right)} = \widetilde{O}\big(d\sqrt{T}\big), \qquad \beta_T \approx \sqrt{d \log T}.$$

The LinUCB proof, now complete

1. **The good event** $\theta^* \in \mathcal{C}_t$ for all t holds with probability $1-\delta$: today's theorem, via the vector supermartingale, Ville, and the Gaussian mixture.
2. **One round**, on the good event: optimism and Cauchy–Schwarz, exactly as in the multi-armed UCB proof,

$$r_t \;\le\; \mathrm{UCB}_t(x_t) - \langle x_t, \theta^* \rangle \;\le\; 2\beta_{t-1}\,\|x_t\|_{V_{t-1}^{-1}}\,.$$

3. **Summing the widths**: the elliptical potential lemma, $\sum_t \min\{1, \|x_t\|^2_{V_{t-1}^{-1}}\} \le 2d\log(1 + TL^2/(d\gamma))$, which is linear algebra about any arm sequence.
4. **Cauchy–Schwarz** glues them, so with probability $1-\delta$,

$$R_T \;\le\; 2\beta_T\sqrt{2dT\log\left(1+\tfrac{TL^2}{d\gamma}\right)} = \widetilde{O}\big(d\sqrt{T}\big), \qquad \beta_T \approx \sqrt{d\log T}.$$

Both $\sqrt{d}$'s are accounted for: one from the mixture (the radius β_T), one from the potential lemma (how many wide directions there can be). Only step 1 was ever in doubt, and it is the only step that needed a martingale.

What we did, and what it cost

- **Fixed design:** $\left\|\hat{\theta} - \theta^*\right\|_V \lesssim \sqrt{d} + \sqrt{\log(1/\delta)}$, by independence, Chernoff in each direction, and a union bound over a net of the sphere.
- **Adaptive design, all t:** $\left\|\hat{\theta}_t - \theta^*\right\|_{\bar{V}_t} \lesssim \sqrt{d \log t} + \sqrt{\log(1/\delta)}$, by a supermartingale in each direction λ, Ville's inequality in place of Chernoff, and a Gaussian mixture over λ in place of the union bound over the net.
- **The price, $\log\det \bar{V}_t \approx d \log t$,** is the continuous analog of $\log|\mathcal{C}_\epsilon| \approx d$: the mixture has to spread its mass over directions *and* scales, because it does not know in advance how large V_t will grow. It is also what makes the bound valid for all t at once.
- **One construction, three bounds.** Lines ($h = \delta_\lambda$), curves for means (h Gaussian, $d = 1$), and ellipsoids for regression (h Gaussian on $\mathbb{R}^d$). The last one is what every linear bandit algorithm with adaptive data rests on: LinUCB's index, Thompson's inflation, and the analysis of both.

Adversarial linear bandits

EXP3 with features, and a design in place of uniform exploration

The setting, and what we already have

Adversarial linear bandit with finitely many arms (notes §7.3)

Known: arms $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$. **Adversary** fixes $\theta_1, \dots, \theta_T \in \mathbb{R}^d$ in advance, $|\langle x_i, \theta_t \rangle| \le 1$.
for $t = 1, \dots, T$: draw $I_t \sim q_t \in \triangle_n$; suffer and observe $\ell_{t,I_t} = \langle x_{I_t}, \theta_t \rangle$ only.
Regret: $\max_i \mathbb{E}\big[\sum_t \langle x_{I_t} - x_i, \theta_t \rangle\big]$.

No noise, no distribution. Ignoring the features, EXP3 gives $\sqrt{12\, nT \log n}$ (Lecture 8). Can the linear structure make d replace n, as in the stochastic case?

The setting, and what we already have

Adversarial linear bandit with finitely many arms (notes §7.3)

Known: arms $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$. **Adversary** fixes $\theta_1, \dots, \theta_T \in \mathbb{R}^d$ in advance, $|\langle x_i, \theta_t \rangle| \le 1$.
for $t = 1, \dots, T$: draw $I_t \sim q_t \in \triangle_n$; suffer and observe $\ell_{t,I_t} = \langle x_{I_t}, \theta_t \rangle$ only.
Regret: $\max_i \mathbb{E}\big[\sum_t \langle x_{I_t} - x_i, \theta_t \rangle\big]$.

No noise, no distribution. Ignoring the features, EXP3 gives $\sqrt{12\, nT \log n}$ (Lecture 8). Can the linear structure make d replace n, as in the stochastic case?

Recall Theorem 12 (Lecture 8)

$I_t \sim q_t = (1-\gamma)p_t + \gamma\lambda$, p_t exponential weights on estimates $\hat{\ell}_{t,i}$ with $\mathbb{E}[\hat{\ell}_{t,i} \mid \mathcal{F}_{t-1}] = \ell_{t,i}$ and $-\eta\hat{\ell}_{t,i} \le 1$:

$$\max_i \mathbb{E}\Big[\sum_{t=1}^{T} \ell_{t,I_t} - \ell_{t,i}\Big] \;\le\; 2\gamma T + \frac{\log n}{\eta} + (1-\gamma)\,\eta\,\mathbb{E}\Big[\sum_{t=1}^{T}\sum_{j=1}^{n} p_{t,j}\hat{\ell}_{t,j}^2\Big].$$

The estimator: least squares with one sample

Write $A(q) = \sum_{i=1}^{n} q_i \, x_i x_i^\top$, the design matrix of the sampling distribution. Estimate the whole loss vector θ_t from the single observation (x_{I_t}, ℓ_{t,I_t}):

$$\hat{\theta}_t \;=\; A(q_t)^{-1} \, x_{I_t} \, \ell_{t,I_t}, \qquad \hat{\ell}_{t,i} \;=\; \langle x_i, \hat{\theta}_t \rangle = x_i^\top A(q_t)^{-1} x_{I_t} \, \ell_{t,I_t}.$$

The estimator: least squares with one sample

Write $A(q) = \sum_{i=1}^{n} q_i \, x_i x_i^\top$, the design matrix of the sampling distribution. Estimate the whole loss vector θ_t from the single observation (x_{I_t}, ℓ_{t,I_t}):

$$\hat{\theta}_t \;=\; A(q_t)^{-1} \, x_{I_t} \, \ell_{t,I_t}, \qquad \hat{\ell}_{t,i} \;=\; \langle x_i, \hat{\theta}_t \rangle = x_i^\top A(q_t)^{-1} x_{I_t} \, \ell_{t,I_t}.$$

Unbiased. Since $I_t \sim q_t$ given the past and $\ell_{t,I_t} = x_{I_t}^\top \theta_t$,

$$\mathbb{E}[\hat{\theta}_t \mid \mathcal{F}_{t-1}] = A(q_t)^{-1} \, \mathbb{E}\big[x_{I_t} x_{I_t}^\top \mid \mathcal{F}_{t-1}\big] \, \theta_t =$$

The estimator: least squares with one sample

Write $A(q) = \sum_{i=1}^{n} q_i \, x_i x_i^\top$, the design matrix of the sampling distribution. Estimate the whole loss vector θ_t from the single observation (x_{I_t}, ℓ_{t,I_t}):

$$\hat{\theta}_t \;=\; A(q_t)^{-1} \, x_{I_t} \, \ell_{t,I_t}, \qquad \hat{\ell}_{t,i} \;=\; \langle x_i, \hat{\theta}_t \rangle = x_i^\top A(q_t)^{-1} x_{I_t} \, \ell_{t,I_t}.$$

Unbiased. Since $I_t \sim q_t$ given the past and $\ell_{t,I_t} = x_{I_t}^\top \theta_t$,

$$\mathbb{E}[\hat{\theta}_t \mid \mathcal{F}_{t-1}] = A(q_t)^{-1} \, \mathbb{E}\big[x_{I_t} x_{I_t}^\top \mid \mathcal{F}_{t-1}\big] \, \theta_t = A(q_t)^{-1} \Big(\sum_{j=1}^{n} q_{t,j} x_j x_j^\top \Big) \theta_t = \theta_t,$$

so $\mathbb{E}[\hat{\ell}_{t,i} \mid \mathcal{F}_{t-1}] = \langle x_i, \theta_t \rangle = \ell_{t,i}$ for *every* arm, including arms never played.

The estimator: least squares with one sample

Write $A(q) = \sum_{i=1}^{n} q_i \, x_i x_i^\top$, the design matrix of the sampling distribution. Estimate the whole loss vector θ_t from the single observation (x_{I_t}, ℓ_{t,I_t}):

$$\hat{\theta}_t \;=\; A(q_t)^{-1} \, x_{I_t} \, \ell_{t,I_t}, \qquad \hat{\ell}_{t,i} \;=\; \langle x_i, \hat{\theta}_t \rangle = x_i^\top A(q_t)^{-1} x_{I_t} \, \ell_{t,I_t}.$$

Unbiased. Since $I_t \sim q_t$ given the past and $\ell_{t,I_t} = x_{I_t}^\top \theta_t$,

$$\mathbb{E}[\hat{\theta}_t \mid \mathcal{F}_{t-1}] = A(q_t)^{-1} \, \mathbb{E}\big[x_{I_t} x_{I_t}^\top \mid \mathcal{F}_{t-1}\big] \, \theta_t = A(q_t)^{-1} \Big(\sum_{j=1}^{n} q_{t,j} x_j x_j^\top \Big) \theta_t = \theta_t,$$

so $\mathbb{E}[\hat{\ell}_{t,i} \mid \mathcal{F}_{t-1}] = \langle x_i, \theta_t \rangle = \ell_{t,i}$ for *every* arm, including arms never played.

- Compare the multi-armed estimator $\hat{\ell}_{t,i} = \mathbf{1}\{I_t = i\} \ell_{t,I_t} / q_{t,i}$: it is the special case $\mathcal{X} = \{e_i\}$, where $A(q) = \mathrm{diag}(q)$. Importance weighting *is* one-sample least squares.
- One pull informs every arm, through $A(q_t)^{-1}$. The price is variance: $\hat{\ell}_{t,i}$ can be as large as $\|x_i\|_{A(q_t)^{-1}} \, \|x_{I_t}\|_{A(q_t)^{-1}}$ (Cauchy–Schwarz), which is where the exploration distribution λ will matter.

Checking the two conditions

Boundedness. $q_t = (1-\gamma)p_t + \gamma\lambda$ gives $A(q_t) \succeq \gamma A(\lambda)$, so $\|x\|^2_{A(q_t)^{-1}} \le \gamma^{-1} \|x\|^2_{A(\lambda)^{-1}}$. Cauchy–Schwarz and $|\ell_{t,I_t}| \le 1$:

$|\eta\hat{\ell}_{t,i}| \le \eta\, |x_i^\top A(q_t)^{-1} x_{I_t}| \; \le$

Checking the two conditions

Boundedness. $q_t = (1-\gamma)p_t + \gamma\lambda$ gives $A(q_t) \succeq \gamma A(\lambda)$, so $\|x\|^2_{A(q_t)^{-1}} \le \gamma^{-1} \|x\|^2_{A(\lambda)^{-1}}$. Cauchy–Schwarz and $|\ell_{t,I_t}| \le 1$:

$$|\eta\hat{\ell}_{t,i}| \le \eta\, |x_i^\top A(q_t)^{-1} x_{I_t}| \;\le\; \eta\, \|x_i\|_{A(q_t)^{-1}} \|x_{I_t}\|_{A(q_t)^{-1}} \;\le\; \tfrac{\eta}{\gamma} \max_j \|x_j\|^2_{A(\lambda)^{-1}}$$

We choose λ: make the right side small, i.e. the G-optimal design, and Kiefer–Wolfowitz says $\max_j \|x_j\|^2_{A(\lambda^\star)^{-1}} = d$. With $\gamma = \eta d$, $|\eta\hat{\ell}_{t,i}| \le 1$.

Checking the two conditions

Boundedness. $q_t = (1-\gamma)p_t + \gamma\lambda$ gives $A(q_t) \succeq \gamma A(\lambda)$, so $\|x\|^2_{A(q_t)^{-1}} \le \gamma^{-1} \|x\|^2_{A(\lambda)^{-1}}$. Cauchy–Schwarz and $|\ell_{t,I_t}| \le 1$:

$$|\eta\hat{\ell}_{t,i}| \le \eta\,|x_i^\top A(q_t)^{-1} x_{I_t}| \;\le\; \eta\,\|x_i\|_{A(q_t)^{-1}}\,\|x_{I_t}\|_{A(q_t)^{-1}} \;\le\; \tfrac{\eta}{\gamma}\max_j \|x_j\|^2_{A(\lambda)^{-1}}$$

We choose λ: make the right side small, i.e. the G-optimal design, and Kiefer–Wolfowitz says $\max_j \|x_j\|^2_{A(\lambda^\star)^{-1}} = d$. With $\gamma = \eta d$, $|\eta\hat{\ell}_{t,i}| \le 1$.

Second moment. With $\ell^2 \le 1$ and $\mathbb{E}[x_{I_t} x_{I_t}^\top \mid \mathcal{F}_{t-1}] = A(q_t) \succeq (1-\gamma)A(p_t)$:

$$\begin{aligned}
\mathbb{E}\Big[\sum_j p_{t,j}\hat{\ell}^2_{t,j} \,\Big|\, \mathcal{F}_{t-1}\Big] &\le \sum_j p_{t,j}\, x_j^\top A(q_t)^{-1}\, \mathbb{E}\big[x_{I_t} x_{I_t}^\top \mid \mathcal{F}_{t-1}\big] A(q_t)^{-1} x_j \\
&= \\
&\le
\end{aligned}$$

Checking the two conditions

Boundedness. $q_t = (1-\gamma)p_t + \gamma\lambda$ gives $A(q_t) \succeq \gamma A(\lambda)$, so $\|x\|^2_{A(q_t)^{-1}} \le \gamma^{-1}\|x\|^2_{A(\lambda)^{-1}}$. Cauchy–Schwarz and $|\ell_{t,I_t}| \le 1$:

$$|\eta\hat{\ell}_{t,i}| \le \eta\,|x_i^\top A(q_t)^{-1}x_{I_t}| \;\le\; \eta\,\|x_i\|_{A(q_t)^{-1}}\,\|x_{I_t}\|_{A(q_t)^{-1}} \;\le\; \tfrac{\eta}{\gamma}\max_j \|x_j\|^2_{A(\lambda)^{-1}}$$

We choose λ: make the right side small, i.e. the G-optimal design, and Kiefer–Wolfowitz says $\max_j \|x_j\|^2_{A(\lambda^\star)^{-1}} = d$. With $\gamma = \eta d$, $|\eta\hat{\ell}_{t,i}| \le 1$.

Second moment. With $\ell^2 \le 1$ and $\mathbb{E}[x_{I_t}x_{I_t}^\top \mid \mathcal{F}_{t-1}] = A(q_t) \succeq (1-\gamma)A(p_t)$:

$$\begin{aligned}
\mathbb{E}\Big[\sum_j p_{t,j}\hat{\ell}^2_{t,j}\,\Big|\,\mathcal{F}_{t-1}\Big] &\le \sum_j p_{t,j}\,x_j^\top A(q_t)^{-1}\,\mathbb{E}\big[x_{I_t}x_{I_t}^\top \mid \mathcal{F}_{t-1}\big]A(q_t)^{-1}x_j \\
&= \sum_j p_{t,j}\,\|x_j\|^2_{A(q_t)^{-1}} \\
&\le \frac{1}{1-\gamma}\,\mathrm{tr}\Big(A(p_t)^{-1}\sum_j p_{t,j}x_jx_j^\top\Big) = \frac{d}{1-\gamma}.
\end{aligned}$$

The trace identity again: the p_t-average leverage is exactly d, whatever p_t is.

EXP3 for adversarial linear bandits: $\sqrt{12\,dT\log n}$

Corollary (notes Cor. 4)

EXP3(γ) with $\hat{\ell}_{t,i} = x_i^\top A(q_t)^{-1} x_{I_t} \ell_{t,I_t}$, λ the G-optimal design for $\mathcal{X}$, $\eta = \sqrt{\log(n)/(3dT)}$, $\gamma = \eta d$:

$$\max_i \; \mathbb{E}\Big[\sum_{t=1}^{T} \langle x_{I_t} - x_i, \theta_t \rangle\Big] \;\le\; \sqrt{12\,dT\log n}.$$

EXP3 for adversarial linear bandits: $\sqrt{12\,dT\log n}$

Corollary (notes Cor. 4)

EXP3(γ) with $\hat{\ell}_{t,i} = x_i^\top A(q_t)^{-1} x_{I_t} \ell_{t,I_t}$, λ the G-optimal design for $\mathcal{X}$, $\eta = \sqrt{\log(n)/(3dT)}$, $\gamma = \eta d$:

$$\max_i \; \mathbb{E}\Big[\sum_{t=1}^{T} \langle x_{I_t} - x_i, \theta_t \rangle\Big] \;\le\; \sqrt{12\,dT\log n}.$$

Proof. Plug the two bounds into Theorem 12:

$$\text{regret} \le 2\gamma T + \frac{\log n}{\eta} + (1-\gamma)\eta \cdot \frac{dT}{1-\gamma} =$$

EXP3 for adversarial linear bandits: $\sqrt{12\,dT\log n}$

Corollary (notes Cor. 4)

EXP3(γ) with $\hat{\ell}_{t,i} = x_i^\top A(q_t)^{-1} x_{I_t} \ell_{t,I_t}$, λ the G-optimal design for $\mathcal{X}$, $\eta = \sqrt{\log(n)/(3dT)}$, $\gamma = \eta d$:

$$\max_i \; \mathbb{E}\Big[\sum_{t=1}^{T} \langle x_{I_t} - x_i, \theta_t \rangle\Big] \;\le\; \sqrt{12\,dT\log n}.$$

Proof. Plug the two bounds into Theorem 12:

$$\text{regret} \le 2\gamma T + \frac{\log n}{\eta} + (1-\gamma)\eta \cdot \frac{dT}{1-\gamma} = 3\eta dT + \frac{\log n}{\eta},$$

balanced by $\eta = \sqrt{\log(n)/(3dT)}$. ■

EXP3 for adversarial linear bandits: $\sqrt{12\,dT\log n}$

Corollary (notes Cor. 4)

EXP3(γ) with $\hat{\ell}_{t,i} = x_i^\top A(q_t)^{-1} x_{I_t} \ell_{t,I_t}$, λ the G-optimal design for $\mathcal{X}$, $\eta = \sqrt{\log(n)/(3dT)}$, $\gamma = \eta d$:

$$\max_i \; \mathbb{E}\Big[\sum_{t=1}^{T} \langle x_{I_t} - x_i, \theta_t \rangle\Big] \;\le\; \sqrt{12\,dT\log n}.$$

Proof. Plug the two bounds into Theorem 12:

$$\text{regret} \le 2\gamma T + \frac{\log n}{\eta} + (1-\gamma)\eta \cdot \frac{dT}{1-\gamma} = 3\eta dT + \frac{\log n}{\eta},$$

balanced by $\eta = \sqrt{\log(n)/(3dT)}$. ■

- d **replaces** n **again;** the arm count survives only as $\log n$, as in the stochastic $\sqrt{dT\log(nT)}$. With $\mathcal{X} = \{e_i\}$ this is Corollary 3 verbatim.
- **No stochastic assumption anywhere,** yet the same structure is exploited: a design spreads a small exploration budget so every direction has leverage at most d.

Linear bandits: the scoreboard

Finite arm set $\mathcal{X} \subset \mathbb{R}^d$, $|\mathcal{X}| = n$, horizon T

setting	algorithm	regret
stochastic	elimination with designs	$\sqrt{dT \log(nT)}$
stochastic, any $\mathcal{X}_t$	LinUCB / linear Thompson	$d\sqrt{T}\log T$ / $d^{3/2}\sqrt{T}$
adversarial	EXP3 with a G-optimal design	$\sqrt{12\, dT \log n}$
lower bound	any algorithm, $\mathcal{X} = \{e_i\}$	$\frac{1}{27}\sqrt{(d-1)T}$

- For a finite arm set, adversarial and stochastic cost the same up to logs, both within logs of the lower bound. Optimism pays an extra $\sqrt{d}$ for arbitrary, changing arm sets.
- Two confidence tools for the stochastic side, both now proved: fixed-design intervals plus a union bound (elimination), and the martingale ellipsoid (LinUCB).
- On the adversarial side the design caps the variance of the one-sample estimator *and* is the exploration distribution: Kiefer–Wolfowitz, a third time.