CSE 541: Interactive Learning

Lecture 11: LinUCB, Thompson sampling, and martingale confidence sets

Kevin Jamieson

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

Wednesday, November 4, 2026

Review, 1: least squares and its error in a direction

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

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

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

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

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

Review, 2: experimental design and elimination

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

Kiefer–Wolfowitz

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

Proposition

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

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

Elimination with designs: rounds of $\tau_\ell = 2d\epsilon_\ell^{-2} \log(4\ell^2 n/\delta)$ pulls from a G-optimal design over the survivors, fresh samples each round so that the fixed-design interval applies; regret $\lesssim \sqrt{dT \log(nT)}$, d in place of n.

Review, 3: the fixed-design confidence ellipsoid, by covering

Fixed design, $V = X^\top X$. The one-direction bound $(*)$: for fixed z, $\langle z, \hat{\theta} - \theta^* \rangle \le \|z\|_{V^{-1}} \sqrt{2 \log(1/\delta)}$ w.p. $1 - \delta$. We wanted all directions at once, $\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2 = 1} \langle V^{1/2} u, \hat{\theta} - \theta^* \rangle$, whose maximizer U is random.

Review, 3: the fixed-design confidence ellipsoid, by covering

Fixed design, $V = X^\top X$. The one-direction bound $(*)$: for fixed z, $\langle z, \hat{\theta} - \theta^* \rangle \le \|z\|_{V^{-1}} \sqrt{2 \log(1/\delta)}$ w.p. $1 - \delta$. We wanted all directions at once, $\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2 = 1} \langle V^{1/2} u, \hat{\theta} - \theta^* \rangle$, whose maximizer U is random.

1. **Cover the sphere** by $\mathcal{C}_\epsilon$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ (volume argument).

Review, 3: the fixed-design confidence ellipsoid, by covering

Fixed design, $V = X^\top X$. The one-direction bound $(*)$: for fixed z, $\langle z, \hat{\theta} - \theta^* \rangle \le \|z\|_{V^{-1}} \sqrt{2\log(1/\delta)}$ w.p. $1-\delta$. We wanted all directions at once, $\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2=1} \langle V^{1/2}u, \hat{\theta} - \theta^* \rangle$, whose maximizer U is random.

1. **Cover the sphere** by $\mathcal{C}_\epsilon$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ (volume argument).
2. **Union bound over the net:** w.p. $1-\delta$, $\langle V^{1/2}y, \hat{\theta} - \theta^* \rangle \le \sqrt{2\log(|\mathcal{C}_\epsilon|/\delta)}$ for all $y \in \mathcal{C}_\epsilon$, using $\left\|V^{1/2}y\right\|_{V^{-1}} = 1$.

Review, 3: the fixed-design confidence ellipsoid, by covering

Fixed design, $V = X^\top X$. The one-direction bound $(*)$: for fixed z, $\langle z, \hat{\theta} - \theta^* \rangle \le \|z\|_{V^{-1}} \sqrt{2\log(1/\delta)}$ w.p. $1-\delta$. We wanted all directions at once, $\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2=1} \langle V^{1/2}u, \hat{\theta} - \theta^* \rangle$, whose maximizer U is random.

1. **Cover the sphere** by $\mathcal{C}_\epsilon$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ (volume argument).
2. **Union bound over the net:** w.p. $1-\delta$, $\langle V^{1/2}y, \hat{\theta} - \theta^* \rangle \le \sqrt{2\log(|\mathcal{C}_\epsilon|/\delta)}$ for all $y \in \mathcal{C}_\epsilon$, using $\left\|V^{1/2}y\right\|_{V^{-1}} = 1$.
3. **Reach every direction from the net:** with $\|U - y\|_2 \le \epsilon$, Cauchy–Schwarz gives $\left\|\hat{\theta} - \theta^*\right\|_V \le \sqrt{2\log(|\mathcal{C}_\epsilon|/\delta)} + \epsilon \left\|\hat{\theta} - \theta^*\right\|_V$; rearrange with $\epsilon = 1/2$.

Review, 3: the fixed-design confidence ellipsoid, by covering

Fixed design, $V = X^\top X$. The one-direction bound $(*)$: for fixed z, $\langle z, \hat{\theta} - \theta^* \rangle \le \|z\|_{V^{-1}} \sqrt{2\log(1/\delta)}$ w.p. $1-\delta$. We wanted all directions at once, $\left\|\hat{\theta} - \theta^*\right\|_V = \max_{\|u\|_2=1} \langle V^{1/2}u, \hat{\theta} - \theta^* \rangle$, whose maximizer U is random.

1. **Cover the sphere** by $\mathcal{C}_\epsilon$ with $|\mathcal{C}_\epsilon| \le (3/\epsilon)^d$ (volume argument).
2. **Union bound over the net:** w.p. $1-\delta$, $\langle V^{1/2}y, \hat{\theta} - \theta^* \rangle \le \sqrt{2\log(|\mathcal{C}_\epsilon|/\delta)}$ for all $y \in \mathcal{C}_\epsilon$, using $\left\|V^{1/2}y\right\|_{V^{-1}} = 1$.
3. **Reach every direction from the net:** with $\|U - y\|_2 \le \epsilon$, Cauchy–Schwarz gives $\left\|\hat{\theta} - \theta^*\right\|_V \le \sqrt{2\log(|\mathcal{C}_\epsilon|/\delta)} + \epsilon \left\|\hat{\theta} - \theta^*\right\|_V$; rearrange with $\epsilon = 1/2$.

Fixed-design confidence ellipsoid

W.p. $\ge 1-\delta$: $\quad \left\|\hat{\theta} - \theta^*\right\|_V \le 2\sqrt{2d\log 6 + 2\log(1/\delta)} =: \beta_{\text{fixed}}$, $\quad$ radius $\approx \sqrt{d}$, every direction at once.

The $\sqrt{d}$ is the union bound over a net of size $e^{O(d)}$. **What breaks for a bandit:** the design depends on the noise, so $(*)$ itself is false. We stated a replacement last time; its proof is the second half of today.

Optimism, continued

LinUCB and linear Thompson sampling

The confidence ellipsoid for adaptive data (stated last time)

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

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

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

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

The confidence ellipsoid for adaptive data (stated last time)

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

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

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

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

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

From the ellipsoid to an index: optimism

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

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

From the ellipsoid to an index: optimism

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

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

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

From the ellipsoid to an index: optimism

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

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

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

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

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

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

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

Regret, step 1: one round costs one width

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

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

Regret, step 1: one round costs one width

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

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

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

Regret, step 1: one round costs one width

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

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

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

What is left

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

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

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

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

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

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

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

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

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

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

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

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

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

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

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

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

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

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

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

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

Regret, step 2: the elliptical potential lemma

Lemma (Lattimore–Szepesvári Lemma 19.4)

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

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

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

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

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

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

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

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

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

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

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

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

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

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

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

$$\le$$

Regret, step 3: Cauchy–Schwarz and done

Theorem (Lattimore–Szepesvári Thm. 19.2)

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

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

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

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

■

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

Linear Thompson sampling

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

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

the same center and shape as the confidence ellipsoid.

Linear Thompson sampling (Agrawal & Goyal 2013)

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

Linear Thompson sampling

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

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

the same center and shape as the confidence ellipsoid.

Linear Thompson sampling (Agrawal & Goyal 2013)

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

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

What is known about linear Thompson sampling

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

Three algorithms

Linear bandits so far

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

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

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 $t \leq T$ and paid a union bound. What is the right price for every $t \in \mathbb{N}$ simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \ldots, y_{t-1}$, so 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 $t \in \mathbb{N}$ simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \dots, y_{t-1}$, so the summands $x_t \eta_t$ are not independent and the fixed-design calculation is wrong.

The setting. A filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ (what is known after round t); Z_t is *adapted* if it is $\mathcal{F}_t$-measurable, *predictable* if $\mathcal{F}_{t-1}$-measurable. In a bandit, x_t is predictable; η_t, y_t are adapted.

Definition

An adapted, integrable M_t is a **martingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] = M_{t-1}$, a **supermartingale** if $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] \le M_{t-1}$: 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 $t \in \mathbb{N}$ simultaneously?
2. **Adaptive designs.** In a bandit x_t depends on $y_1, \dots, y_{t-1}$, so the summands $x_t \eta_t$ are not independent and the fixed-design calculation is wrong.

The setting. A filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ (what is known after round t); Z_t is *adapted* if it is $\mathcal{F}_t$-measurable, *predictable* if $\mathcal{F}_{t-1}$-measurable. In a bandit, x_t is predictable; η_t, y_t are adapted.

Definition

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

Examples: a sum of conditionally mean-zero increments is a martingale; a gambler's fortune in a fair casino is a martingale, in a real one a supermartingale.

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

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

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

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

Peel off the last increment:

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

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

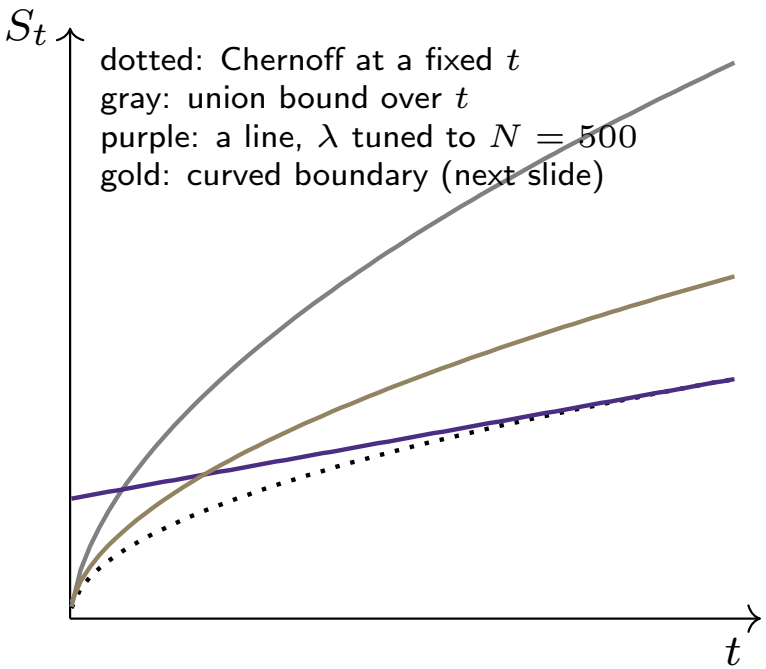
Apply Ville to $M_t(\lambda) = e^{\lambda S_t - t\lambda^2/2}$ for a fixed $\lambda > 0$:
$M_t(\lambda) \geq 1/\delta \iff S_t \geq \frac{t\lambda}{2} + \frac{\log(1/\delta)}{\lambda}$,

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

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 $|S_t|$:

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

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.