INVENTEURS DU MONDE NUMÉRIQUE

Spectral Bandits for Smooth Graph Functions

Tomáš Kocák — *INRIA Lille - Nord Europe, France*
Michal Valko — *INRIA Lille - Nord Europe, France*
Rémi Munos — *INRIA Lille - Nord Europe & Microsoft Research NE*
Branislav Kveton — *Technicolor, Palo Alto*

SequeL – INRIA Lille

ICML 2014

Beijing, June 2014

Movie recommendation: (in each time step)

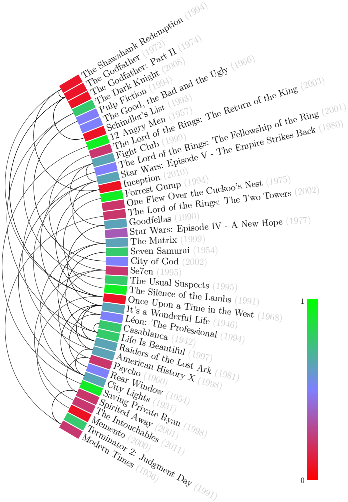
- Recommend movies to a **single user**.
- Good prediction after a few steps ($T \ll N$).

Goal:

- Maximize overall reward (sum of ratings).

Assumptions:

- Unknown reward function $f : V(G) \to \mathbb{R}$.
- Function f is **smooth** on a graph.
- Neighboring movies $\Rightarrow$ similar preferences.
- Similar preferences $\not\Rightarrow$ neighboring movies.



Smooth graph function

- Graph G with vertex set $V(G) = \{1, \ldots, N\}$ and edge set $E(G)$.
- $f_1, \ldots, f_N$: Values of the function on the vertices of the graph.
- $w_{i,j}$: Weight of the edge connecting nodes i and j.
- **Smoothness of the function:**

$$S_G(f) = \frac{1}{2} \sum_{i,j \leq N} w_{i,j} (f_i - f_j)^2$$

  - Smaller value of $S_G(f)$, smoother the function f is.

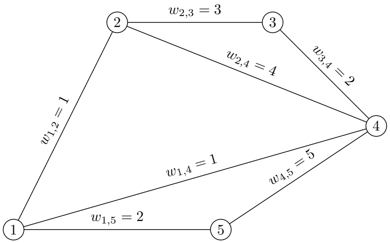
- **Examples:**
  - **Complete graph:** Only constant function has smoothness 0.
  - **Edgeless graph:** Every function has smoothness 0.
  - **Constant function:** Smoothness 0 for every graph.

Graph Laplacian

- $\mathcal{W}$: $N \times N$ matrix of the edge weights $w_{i,j}$.
- $\mathcal{D}$: Diagonal matrix with the entries $d_i = \sum_j w_{i,j}$.
- $\mathcal{L} = \mathcal{D} - \mathcal{W}$: Graph Laplacian.
  - Positive semidefinite matrix.
  - Diagonally dominant matrix.

Example:

$$\mathcal{L} = \begin{pmatrix} 4 & -1 & 0 & -1 & -2 \\ -1 & 8 & -3 & -4 & 0 \\ 0 & -3 & 5 & -2 & 0 \\ -1 & -4 & -2 & 12 & -5 \\ -2 & 0 & 0 & -5 & 7 \end{pmatrix}$$



Smoothness of the function and Laplacian

- $\boldsymbol{f} = (f_1, \ldots, f_N)^\top$: Vector of function values.
- Let $\mathcal{L} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top$ be the eigendecomposition of the Laplacian.
  - Diagonal matrix $\boldsymbol{\Lambda}$ whose diagonal entries are eigenvalues of $\mathcal{L}$.
  - Columns of $\mathbf{Q}$ are eigenvectors of $\mathcal{L}$.
  - Columns of $\mathbf{Q}$ form a basis.
- $\boldsymbol{\alpha}^*$: Unique vector such that $\mathbf{Q}\boldsymbol{\alpha}^* = \boldsymbol{f}$ Note: $\mathbf{Q}^\top\boldsymbol{f} = \boldsymbol{\alpha}^*$

$$S_G(f) = \boldsymbol{f}^\top\mathcal{L}\boldsymbol{f} = \boldsymbol{f}^\top\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top\boldsymbol{f} = \boldsymbol{\alpha}^{*\top}\boldsymbol{\Lambda}\boldsymbol{\alpha}^* = \|\boldsymbol{\alpha}^*\|_{\boldsymbol{\Lambda}} = \sum_{i=1}^{N}\lambda_i(\alpha_i^*)^2$$

Smoothness and regularization: Small value of

(a) $S_G(f)$ **(b)** $\boldsymbol{\Lambda}$ norm of $\boldsymbol{\alpha}^*$ **(c)** α_i^* for large λ_i

Setting

Problem structure

- Underlying graph structure encoded in the graph laplacian $\mathcal{L}$.
- Eigendecomposition of graph laplacian $\mathcal{L} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\mathsf{T}$ where $\mathbf{Q}$ is the matrix with eigenvectors in columns.
- the i-th row $\mathbf{x}_i$ of the matrix $\mathbf{Q}$ corresponds to the arm i.

Learning setting

- In each time step choose a node $\pi(t)$.
- Obtain noisy reward $r_t = \mathbf{x}_{\pi(t)}^\mathsf{T}\boldsymbol{\alpha}^* + \varepsilon_t$.   Note: $\mathbf{x}_{\pi(t)}^\mathsf{T}\boldsymbol{\alpha}^* = f_{\pi(t)}$
  - ε_t is R-sub-Gaussian noise. $\forall \xi \in \mathbb{R}, \mathbb{E}[e^{\xi\varepsilon_t}] \leq \exp\left(\xi^2 R^2/2\right)$
- Minimize cumulative regret

$$R_T = T \max_a \left(\mathbf{x}_a^\mathsf{T}\boldsymbol{\alpha}^*\right) - \sum_{t=1}^{T} \mathbf{x}_{\pi(t)}^\mathsf{T}\boldsymbol{\alpha}^*.$$

Solutions

- **Linear bandit algorithms** (Existing solutions)
  - **LinUCB** (Li et al., 2010)
    - Regret bound $\approx D\sqrt{T \ln T}$
  - **SupLinRel** (Auer, 2002)
    - Regret bound $\approx \sqrt{DT \ln T}$

Note: D is ambient dimension, in our case N, length of x_i.

Solutions

- **Linear bandit algorithms** (Existing solutions)
  - **LinUCB** (Li et al., 2010)
    - Regret bound $\approx D\sqrt{T \ln T}$
  - **SupLinRel** (Auer, 2002)
    - Regret bound $\approx \sqrt{DT \ln T}$

Note: D is ambient dimension, in our case N, length of x_i.

- **Spectral bandit algorithms** (Our solutions)
  - **SpectralUCB**
    - Regret bound $\approx d\sqrt{T \ln T}$
  - **SpectralEliminator**
    - Regret bound $\approx \sqrt{dT \ln T}$

Note: d is effective dimension, usually much smaller than D.

Effective dimension

- **Effective dimension:** Largest d such that

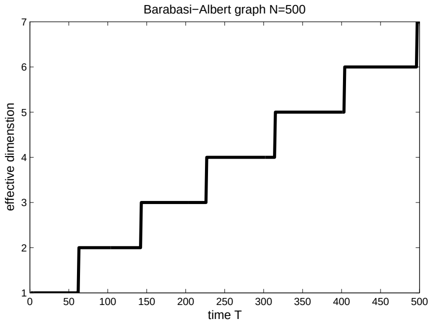
$$(d-1)\lambda_d \leq \frac{T}{\ln(1+T/\lambda)}.$$

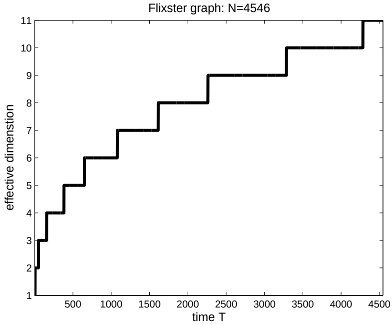
- λ_i : i-th smallest eigenvalue of $\mathbf{\Lambda}$.
- λ: Regularization parameter of the algorithm.

Properties:

- d is small when the coefficients λ_i grow rapidly above time.
- d is related to the number of "non-negligible" dimensions.
- Usually d is much smaller than D in real world graphs.
- Can be computed beforehand.

Effective dimension vs. Ambient dimension





$$d \ll D$$

Note: In our setting $T < N = D$.

SpectralUCB algorithm

Input:
N, T, $\{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\}$, λ, δ, R, C
Run:
$\mathbf{\Lambda} \leftarrow \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
$d \leftarrow \max\{d : (d-1)\lambda_d \leq T / \ln(1 + T/\lambda)\}$
for $t = 1$ **to** T **do**
    Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:
        $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^\mathsf{T}$
        $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^\mathsf{T}$
        $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^\mathsf{T} + \mathbf{\Lambda}$
        $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^\mathsf{T} \mathbf{r}$
        $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$
        $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^\mathsf{T} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}} \right)$
        Observe the reward r_t
end for

SpectralUCB algorithm

Input:

N, T, $\{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\}$, λ, δ, R, C

Run:

$\mathbf{\Lambda} \leftarrow \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$

$d \leftarrow \max\{d : (d-1)\lambda_d \leq T / \ln(1 + T/\lambda)\}$

for $t = 1$ **to** T **do**

  Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:

    $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^{\mathsf{T}}$

    $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^{\mathsf{T}}$

    $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^{\mathsf{T}} + \mathbf{\Lambda}$

    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^{\mathsf{T}} \mathbf{r}$

    $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$

    $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^{\mathsf{T}} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}} \right)$

    Observe the reward r_t

end for

SpectralUCB algorithm

Input:
$N,\ T,\ \{\boldsymbol{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \delta,\ R,\ C$
Run:
$\boldsymbol{\Lambda} \leftarrow \boldsymbol{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
$d \leftarrow \max\{d : (d-1)\lambda_d \leq T/\ln(1 + T/\lambda)\}$
for $t = 1$ **to** T **do**
    Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:
        $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^{\mathsf{T}}$
        $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^{\mathsf{T}}$
        $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^{\mathsf{T}} + \boldsymbol{\Lambda}$
        $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^{\mathsf{T}} \mathbf{r}$
        $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$
        $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^{\mathsf{T}} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}} \right)$
        Observe the reward r_t
end for

SpectralUCB algorithm

Input:

N, T, $\{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\}$, λ, δ, R, C

Run:

$\mathbf{\Lambda} \leftarrow \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$

$d \leftarrow \max\{d : (d-1)\lambda_d \leq T / \ln(1 + T/\lambda)\}$

for $t = 1$ **to** T **do**

  Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:

    $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^{\mathsf{T}}$

    $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^{\mathsf{T}}$

    $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^{\mathsf{T}} + \mathbf{\Lambda}$

    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^{\mathsf{T}} \mathbf{r}$

    $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$

    $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^{\mathsf{T}} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}} \right)$

    Observe the reward r_t

end for

SpectralUCB algorithm

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \delta,\ R,\ C$
Run:
$\mathbf{\Lambda} \leftarrow \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
$d \leftarrow \max\{d : (d-1)\lambda_d \leq T/\ln(1 + T/\lambda)\}$
for $t = 1$ **to** T **do**
  Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:
    $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^\mathsf{T}$
    $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^\mathsf{T}$
    $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^\mathsf{T} + \mathbf{\Lambda}$
    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^\mathsf{T} \mathbf{r}$
    $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$
    $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^\mathsf{T} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}} \right)$
    Observe the reward r_t
end for

SpectralUCB algorithm

Input:

N, T, $\{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\}$, λ, δ, R, C

Run:

$\mathbf{\Lambda} \leftarrow \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$

$d \leftarrow \max\{d : (d-1)\lambda_d \leq T / \ln(1 + T/\lambda)\}$

for $t = 1$ **to** T **do**

  Update the basis coefficients $\hat{\boldsymbol{\alpha}}$:

    $\mathbf{X}_t \leftarrow [\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(t-1)}]^\mathsf{T}$

    $\mathbf{r} \leftarrow [r_1, \ldots, r_{t-1}]^\mathsf{T}$

    $\mathbf{V}_t \leftarrow \mathbf{X}_t \mathbf{X}_t^\mathsf{T} + \mathbf{\Lambda}$

    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} \mathbf{X}_t^\mathsf{T} \mathbf{r}$

    $c_t \leftarrow 2R\sqrt{d \ln(1 + t/\lambda) + 2\ln(1/\delta)} + C$

    $\pi(t) \leftarrow \arg\max_a \left(\mathbf{x}_a^\mathsf{T} \hat{\boldsymbol{\alpha}} + c_t \|\mathbf{x}_a\|_{\mathbf{V}_t^{-1}}\right)$

    Observe the reward r_t

end for

SpectralUCB regret bound

- d: Effective dimension.
- λ: Minimal eigenvalue of $\mathbf{\Lambda} = \mathbf{\Lambda}_{\mathcal{L}} + \lambda\mathbf{I}$.
- C: Smoothness upper bound, $\|\boldsymbol{\alpha}^*\|_{\mathbf{\Lambda}} \leq C$.
- $\mathbf{x}_i^\mathsf{T}\boldsymbol{\alpha}^* \in [-1, 1]$ for all i.

The **cumulative regret** R_T of **SpectralUCB** is with probability $1 - \delta$ bounded as

$$R_T \leq \left(8R\sqrt{d \ln \frac{\lambda + T}{\lambda} + 2 \ln \frac{1}{\delta}} + 4C + 4\right)\sqrt{dT \ln \frac{\lambda + T}{\lambda}}.$$

$$R_T \approx d\sqrt{T \ln T}$$

SpectralUCB analysis sketch

- Derivation of the confidence ellipsoid for $\hat{\boldsymbol{\alpha}}$ with probability $1 - \delta$.
  - Using analysis of OFUL (Abbasi-Yadkori et al., 2011)

$$|x^{\mathsf{T}}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^*)| \leq \|x\|_{\mathbf{V}_t^{-1}} \left(R\sqrt{2\ln\left(\frac{|V_t|^{1/2}}{\delta|\boldsymbol{\Lambda}|^{1/2}}\right)} + C \right)$$

- Regret in one time step: $r_t = \mathbf{x}_*^{\mathsf{T}}\boldsymbol{\alpha}^* - \mathbf{x}_{\pi(t)}^{\mathsf{T}}\boldsymbol{\alpha}^* \leq 2c_t\|\mathbf{x}_{\pi(t)}\|_{\mathbf{V}_t^{-1}}$
- Cumulative regret:

$$R_T = \sum_{t=1}^{T} r_t \leq \sqrt{T\sum_{t=1}^{T} r_t^2} \leq 2(\; c_T + 1)\sqrt{2T\ln\frac{|\mathbf{V}_T|}{|\boldsymbol{\Lambda}|}}$$

SpectralUCB analysis sketch

- Derivation of the confidence ellipsoid for $\hat{\boldsymbol{\alpha}}$ with probability $1-\delta$.
  - Using analysis of OFUL (Abbasi-Yadkori et al., 2011)

$$|x^{\mathsf{T}}(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^*)| \leq \|x\|_{\mathbf{V}_t^{-1}} \left(R\sqrt{2\ln\left(\frac{|V_t|^{1/2}}{\delta|\boldsymbol{\Lambda}|^{1/2}}\right)} + C \right)$$

- Regret in one time step: $r_t = \mathbf{x}_*^{\mathsf{T}}\boldsymbol{\alpha}^* - \mathbf{x}_{\pi(t)}^{\mathsf{T}}\boldsymbol{\alpha}^* \leq 2c_t\|\mathbf{x}_{\pi(t)}\|_{\mathbf{V}_t^{-1}}$
- Cumulative regret:

$$R_T = \sum_{t=1}^{T} r_t \leq \sqrt{T\sum_{t=1}^{T} r_t^2} \leq 2(c_T+1)\sqrt{2T\ln\frac{|\mathbf{V}_T|}{|\boldsymbol{\Lambda}|}}$$

SpectralUCB analysis sketch

- Derivation of the confidence ellipsoid for $\hat{\boldsymbol{\alpha}}$ with probability $1-\delta$.
  - Using analysis of OFUL (Abbasi-Yadkori et al., 2011)

$$|x^\mathsf{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^*)| \leq \|x\|_{\mathbf{V}_t^{-1}} \left(R\sqrt{2\ln\left(\frac{|V_t|^{1/2}}{\delta|\boldsymbol{\Lambda}|^{1/2}}\right)} + C \right)$$

- Regret in one time step: $r_t = \mathbf{x}_*^\mathsf{T}\boldsymbol{\alpha}^* - \mathbf{x}_{\pi(t)}^\mathsf{T}\boldsymbol{\alpha}^* \leq 2c_t\|\mathbf{x}_{\pi(t)}\|_{\mathbf{V}_t^{-1}}$
- Cumulative regret:

$$R_T = \sum_{t=1}^T r_t \leq \sqrt{T\sum_{t=1}^T r_t^2} \leq 2(c_T + 1)\sqrt{2T\ln\frac{|\mathbf{V}_T|}{|\boldsymbol{\Lambda}|}}$$

- Upperbound for $\ln(|\mathbf{V}_t|/|\boldsymbol{\Lambda}|)$

$$\ln\frac{|\mathbf{V}_t|}{|\boldsymbol{\Lambda}|} \leq \ln\frac{|\mathbf{V}_T|}{|\boldsymbol{\Lambda}|} \leq 2d\ln\left(\frac{\lambda+T}{\lambda}\right)$$

SpectralEliminator

Input:

$N,\ T,\ \{\boldsymbol{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$

Run:

$A_1 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$

for $j = 1$ **to** J **do**

  $\mathbf{V}_{t_j} \leftarrow \gamma \boldsymbol{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$

  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**

    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:

      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$

    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^\mathsf{T}$

  end for

  Eliminate the arms that are not promising:

    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1}[\mathbf{x}_{t_j}, \ldots, \mathbf{x}_t][r_{t_j}, \ldots, r_t]^\mathsf{T}$

    $A_{j+1} \leftarrow \left\{\mathbf{x} \in A_j, \langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta\right]\right\}$

end for

SpectralEliminator

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
  $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
  end for
  Eliminate the arms that are not promising:
    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1}[\mathbf{x}_{t_j}, \ldots, \mathbf{x}_t][r_{t_j}, \ldots, r_t]^{\mathsf{T}}$
    $A_{j+1} \leftarrow \left\{\mathbf{x} \in A_j, \langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta\right]\right\}$
end for

SpectralEliminator

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
  $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
  end for
  Eliminate the arms that are not promising:
    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} [\mathbf{x}_{t_j}, \ldots, \mathbf{x}_t][r_{t_j}, \ldots, r_t]^{\mathsf{T}}$
    $A_{j+1} \leftarrow \left\{ \mathbf{x} \in A_j, \langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \right] \right\}$
end for

SpectralEliminator

Input:
N, T, $\{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\}$, λ, β, $\{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
    $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
    for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
        Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
            $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
        $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
    end for
    Eliminate the arms that are not promising:
        $\hat{\alpha}_t \leftarrow \mathbf{V}_t^{-1} [\mathbf{x}_{t_j}, \dots, \mathbf{x}_t][r_{t_j}, \dots, r_t]^{\mathsf{T}}$
        $A_{j+1} \leftarrow \left\{ \mathbf{x} \in A_j, \langle \hat{\alpha}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\alpha}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \right] \right\}$
end for

SpectralEliminator

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
  $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
  end for
  Eliminate the arms that are not promising:
    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} [\mathbf{x}_{t_j}, \ldots, \mathbf{x}_t][r_{t_j}, \ldots, r_t]^{\mathsf{T}}$
    $A_{j+1} \leftarrow \left\{ \mathbf{x} \in A_j, \langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \right] \right\}$
end for

SpectralEliminator

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
  $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
  end for
  Eliminate the arms that are not promising:
    $\hat{\boldsymbol{\alpha}}_t \leftarrow \mathbf{V}_t^{-1} [\mathbf{x}_{t_j}, \dots, \mathbf{x}_t][r_{t_j}, \dots, r_t]^{\mathsf{T}}$
    $A_{j+1} \leftarrow \left\{ \mathbf{x} \in A_j, \langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\boldsymbol{\alpha}}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \right] \right\}$
end for

SpectralEliminator

Input:
$N,\ T,\ \{\mathbf{\Lambda}_{\mathcal{L}}, \mathbf{Q}\},\ \lambda,\ \beta,\ \{t_j\}_j^J$
Run:
$A_1 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$
for $j = 1$ **to** J **do**
  $\mathbf{V}_{t_j} \leftarrow \gamma \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$
  for $t = t_j$ **to** $\min(t_{j+1} - 1, T)$ **do**
    Play $\mathbf{x}_t \in A_j$ with the largest width to observe r_t:
      $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in A_j} \|\mathbf{x}\|_{\mathbf{V}_t^{-1}}$
    $\mathbf{V}_{t+1} \leftarrow \mathbf{V}_t + \mathbf{x}_t \mathbf{x}_t^{\mathsf{T}}$
  end for
  Eliminate the arms that are not promising:
    $\hat{\alpha}_t \leftarrow \mathbf{V}_t^{-1} [\mathbf{x}_{t_j}, \ldots, \mathbf{x}_t][r_{t_j}, \ldots, r_t]^{\mathsf{T}}$
    $A_{j+1} \leftarrow \left\{ \mathbf{x} \in A_j, \langle \hat{\alpha}_t, \mathbf{x} \rangle + \|\mathbf{x}\|_{V_t^{-1}} \beta \geq \max_{\mathbf{x} \in A_j} \left[\langle \hat{\alpha}_t, \mathbf{x} \rangle - \|\mathbf{x}\|_{\mathbf{V}_t^{-1}} \beta \right] \right\}$
end for

SpectralEliminator regret bound

- d: Effective dimension.
- λ: Minimal eigenvalue of $\mathbf{\Lambda} = \mathbf{\Lambda}_{\mathcal{L}} + \lambda \mathbf{I}$.
- C: Smoothness upper bound, $\|\boldsymbol{\alpha}^*\|_{\mathbf{\Lambda}} \leq C$.
- $t_j = 2^{j-1}$: Beginning of the phase j.
- $\mathbf{x}_i^\mathsf{T} \boldsymbol{\alpha}^* \in [-1, 1]$ for all i.
- $\beta = 2R\sqrt{14 \ln(2N \log_2 T/\delta)} + C$: Parameter of the elimination.

The **cumulative regret** R_T of **SpectralEliminator** is with probability $1 - \delta$ bounded as

$$R_T \leq \frac{4}{\ln 2}\left(2R\sqrt{14 \ln \frac{2K \log_2 T}{\delta}} + C\right)\sqrt{dT \ln\left(1 + \frac{T}{\lambda}\right)}.$$

$$R_T \approx \sqrt{dT \ln T}$$

SpectralEliminator analysis sketch

- Derivation of the confidence ellipsoid for $\hat{\boldsymbol{\alpha}}$ with probability $1-\delta$.
  - Using analysis of OFUL (Abbasi-Yadkori et al., 2011)

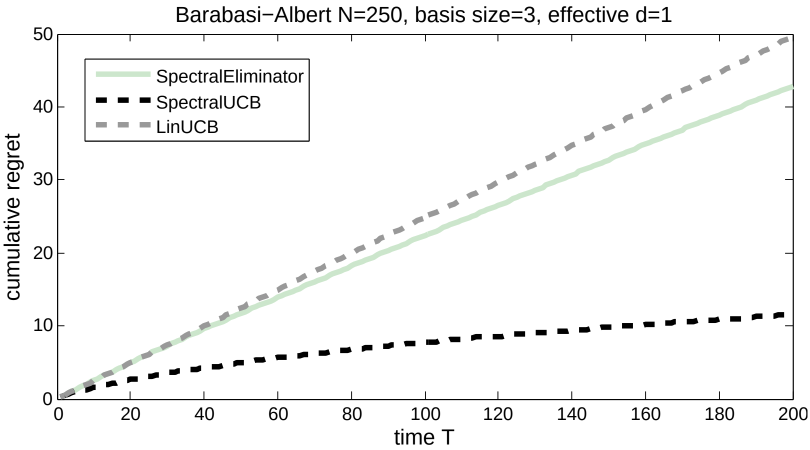
$$|x^{\mathsf{T}}(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^*)| \le \|x\|_{\mathbf{V}_t^{-1}} \left(R\sqrt{2\ln\left(\frac{|V_t|^{1/2}}{\delta|\boldsymbol{\Lambda}|^{1/2}}\right)} + C \right)$$

- Using Azuma-Hoeffding inequality Note: phases are independent

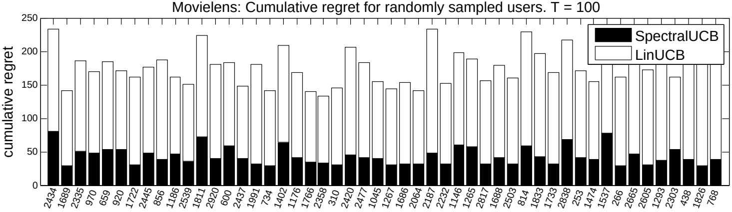
$$R_T \le \sum_{j=0}^{J}(t_{j+1}-t_j)\left[\langle \mathbf{x}^* - \mathbf{x}_t, \hat{\boldsymbol{\alpha}}_j\rangle + (\|\mathbf{x}^*\|_{\mathbf{V}_j^{-1}} + \|\mathbf{x}_t\|_{\mathbf{V}_j^{-1}})\beta\right]$$

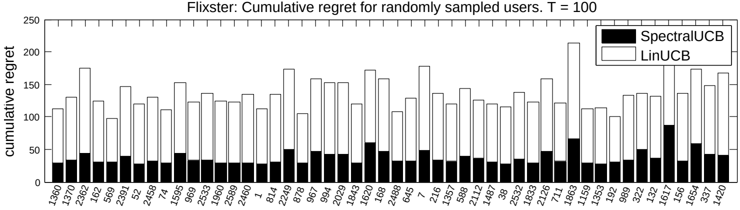
- Bound $\langle \mathbf{x}^* - \mathbf{x}_t, \hat{\boldsymbol{\alpha}}_j\rangle$ for each phase
- No bad arms: $\langle \mathbf{x}^* - \mathbf{x}_t, \hat{\boldsymbol{\alpha}}_j\rangle \le (\|\mathbf{x}^*\|_{\mathbf{V}_j^{-1}} + \|\mathbf{x}_t\|_{\mathbf{V}_j^{-1}})\beta$
- By algorithm: $\|\mathbf{x}\|^2_{\mathbf{V}_j^{-1}} \le \frac{1}{t_j - t_{j-1}} \sum_{s=t_{j-1}+1}^{t_j} \|\mathbf{x}_s\|^2_{\mathbf{V}_{s-1}^{-1}}$

Synthetic experiment



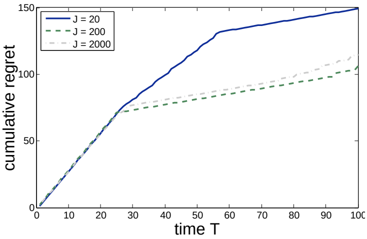
Real world experiment

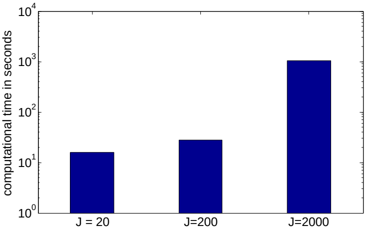




Improving the running time: reduced eigenbasis

- **Reduced basis:** We only need first few eigenvectors.
- **Getting J eigenvectors:** $\mathcal{O}(Jm \log m)$ time for m edges
- Computationally less expensive, comparable performance.





Conclusion

- New spectral bandit setting **(for smooth graph functions)**.
- **SpectralUCB**.
  - Regret bound $\approx d\sqrt{T \ln T}$
- **SpectralEliminator**
  - Regret bound $\approx \sqrt{dT \ln T}$
  - Side result: **LinearEliminator** with $\mathcal{O}(\sqrt{DT \ln T})$ regret for (contextual) linear bandits.
- Bounds scale with **effective dimension** $d \ll D$.
- **SpectralTS** (Thompson Sampling) – AAAI 2014
  - Regret bound $\approx d\sqrt{T \ln N}$
  - Computationally more efficient.

Thank you!

Poster (T8)

Tomáš Kocák
tomas.kocak@inria.fr
sequel.lille.inria.fr

Sylvester's determinant theorem:

$$|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}||\mathbf{I} + \mathbf{A}^{-1}\mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}|(1 + \mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x})$$

Goal:

- Upperbound determinant $|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}|$ for $\|\mathbf{x}\|_2 \leq 1$
- Upperbound $\mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x}$

Sylvester's determinant theorem:

$$|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}||\mathbf{I} + \mathbf{A}^{-1}\mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}|(1 + \mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x})$$

Goal:

- Upperbound determinant $|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}|$ for $\|\mathbf{x}\|_2 \leq 1$
- Upperbound $\mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x}$

$$\mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{Q}\boldsymbol{\Lambda}^{-1}\mathbf{Q}^\mathsf{T}\mathbf{x} = \mathbf{y}^\mathsf{T}\boldsymbol{\Lambda}^{-1}\mathbf{y} = \sum_{i=1}^{N} \lambda_i y_i^2$$

Sylvester's determinant theorem:

$$|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}||\mathbf{I} + \mathbf{A}^{-1}\mathbf{x}\mathbf{x}^\mathsf{T}| = |\mathbf{A}|(1 + \mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x})$$

Goal:

- Upperbound determinant $|\mathbf{A} + \mathbf{x}\mathbf{x}^\mathsf{T}|$ for $\|\mathbf{x}\|_2 \leq 1$
- Upperbound $\mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x}$

$$\mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{Q}^\mathsf{T}\mathbf{x} = \mathbf{y}^\mathsf{T}\mathbf{\Lambda}^{-1}\mathbf{y} = \sum_{i=1}^{N} \lambda_i y_i^2$$

- $\|\mathbf{y}\|_2 \leq 1$.
- $\mathbf{y}$ is a canonical vector.
- $\mathbf{x} = \mathbf{Q}\mathbf{y}$ is an eigenvector of $\mathbf{A}$.

Corollary:
Determinant $|\mathbf{V}_T|$ of $\mathbf{V}_T = \mathbf{\Lambda} + \sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t^\mathsf{T}$ is maximized when all $\mathbf{x}_t$ are aligned with axes.

$$|\mathbf{V}_T| \leq \max_{\sum t_i = T} \prod (\lambda_i + t_i)$$

$$\ln \frac{|\mathbf{V}_T|}{|\mathbf{\Lambda}|} \leq \max_{\sum t_i = T} \sum \ln\left(1 + \frac{t_i}{\lambda_i}\right)$$

$$\ln \frac{|\mathbf{V}_T|}{|\mathbf{\Lambda}|} \leq \sum_{i=1}^{d} \ln\left(1 + \frac{T}{\lambda}\right) + \sum_{i=d+1}^{N} \ln\left(1 + \frac{t_i}{\lambda_{d+1}}\right)$$

$$\leq d \ln\left(1 + \frac{T}{\lambda}\right) + \frac{T}{\lambda_{d+1}}$$

$$\leq 2d \ln\left(1 + \frac{T}{\lambda}\right)$$

$$\boldsymbol{f}^{\mathsf{T}}\mathcal{L}\boldsymbol{f} = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2 = S_G(f)$$

Proof:

$$\boldsymbol{f}^{\mathsf{T}}\mathcal{L}\boldsymbol{f} = \boldsymbol{f}^{\mathsf{T}}\mathcal{D}\boldsymbol{f} - \boldsymbol{f}^{\mathsf{T}}\mathcal{W}\boldsymbol{f} = \sum_{i=1}^{N} d_i f_i^2 - \sum_{i,j\leq N} w_{i,j} f_i f_j$$

$$= \frac{1}{2}\left(\sum_{i=1}^{N} d_i f_i^2 - 2\sum_{i,j\leq N} w_{i,j} f_i f_j + \sum_{j=1}^{N} d_i f_j^2\right) = \frac{1}{2}\sum_{i,j\leq N} w_{i,j}(f_i - f_j)^2$$