Gram-Schmidt walk algorithm with applications to Kómlos conjecture

Witold Bednorz and Piotr Godlewski

Institute of Mathematics
University of Warsaw

HDP conference Bedlewo, 16.06.2023

Kómlos conjecture

- Consider $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$.

Kómlos conjecture

- Consider $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$.
- Suppose that $\|v_i\|_2 \leqslant 1$ for $i \in [n]$.

Kómlos conjecture

- Consider $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$.
- Suppose that $\|v_i\|_2 \leqslant 1$ for $i \in [n]$.
- Let $\varepsilon_i \in \{-1, 1\}$ for $i \in [n]$.

Kómlos conjecture

- Consider $v_1, v_2, \dots, v_n \in \mathbb{R}^d$.
- Suppose that $\|v_i\|_2 \leqslant 1$ for $i \in [n]$.
- Let $\varepsilon_i \in \{-1, 1\}$ for $i \in [n]$.
- For each choice of $(\varepsilon_i)_{i=1}^n$ we define

$$\text{discrepancy} \quad \|\sum_{i=1}^n \varepsilon_i v_i\|_\infty .$$

Kómlos conjecture

- Consider $v_1, v_2, \ldots, v_n \in \mathbb{R}^d$.
- Suppose that $\|v_i\|_2 \leqslant 1$ for $i \in [n]$.
- Let $\varepsilon_i \in \{-1, 1\}$ for $i \in [n]$.
- For each choice of $(\varepsilon_i)_{i=1}^n$ we define

$$\text{discrepancy} \ \ \|\sum_{i=1}^{n} \varepsilon_i v_i\|_\infty.$$

- **Kómlos Conjecture** One can choose $(\varepsilon_i)_{i=1}^n$ such that

$$\|\sum_{i=1}^{n} \varepsilon_i v_i\|_\infty \leqslant C.$$

Some history

Some history

- **Theorem**
  Spencer 1985 There exists $(\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n$, such that

$$\|\sum_{i=1}^{n} \varepsilon_i v_i\|_2 \leqslant C\sqrt{n}\sqrt{1 + \log\left(\frac{d}{n}\right)}.$$

Some history

- **Theorem**
  Spencer 1985 There exists $(\varepsilon_i)_{i=1}^n \in \{-1,1\}^n$*, such that*

$$\|\sum_{i=1}^n \varepsilon_i v_i\|_2 \leqslant C\sqrt{n}\sqrt{1+\log\left(\frac{d}{n}\right)}.$$

  - Beck, Spencer 1980 Pigeonhole + Probabilistic Method.

Some history

- **Theorem**
  Spencer 1985 *There exists* $(\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n$*, such that*

$$\|\sum_{i=1}^{n} \varepsilon_i v_i\|_2 \leqslant C\sqrt{n}\sqrt{1 + \log\left(\frac{d}{n}\right)}.$$

  - Beck, Spencer 1980 Pigeonhole + Probabilistic Method.
  - Gluskin 1987 Convex Geometry approach.

Some history

- **Theorem**
  Spencer 1985 There exists $(\varepsilon_i)_{i=1}^n \in \{-1,1\}^n$*, such that*

$$\|\sum_{i=1}^n \varepsilon_i v_i\|_2 \leqslant C\sqrt{n}\sqrt{1+\log\left(\frac{d}{n}\right)}.$$

  - Beck, Spencer 1980 Pigeonhole + Probabilistic Method.
  - Gluskin 1987 Convex Geometry approach.

Some history

- **Theorem**
  Spencer 1985 There exists $(\varepsilon_i)_{i=1}^n \in \{-1,1\}^n$, *such that*

$$\|\sum_{i=1}^n \varepsilon_i v_i\|_2 \leqslant C\sqrt{n}\sqrt{1+\log\left(\frac{d}{n}\right)}.$$

  - Beck, Spencer 1980 Pigeonhole + Probabilistic Method.
  - Gluskin 1987 Convex Geometry approach.

- **Theorem**
  Banaszczyk, 1998 There exists $(\varepsilon_i)_{i=1}^n \in \{\pm 1\}_{i=1}^n$ *such that*

$$\|\sum_{i=1}^n \varepsilon_i v_i\|_\infty \leqslant C\sqrt{1+\log d}.$$

Banaszczyk's Theorem by GSW algorithm

- Let $M = (v_1, \ldots, v_n)$, M-$d \times n$ matrix.

Banaszczyk's Theorem by GSW algorithm

- Let $M = (v_1, \ldots, v_n)$, M-$d \times n$ matrix.
- GSW algorithm: there exists a r.v. $X \in \{-1, 1\}^n \subset \mathbb{R}^n$ s.t.

$$\mathbf{E} \exp\left(\langle MX, v\rangle\right) \leqslant \exp\left(\frac{1}{2}\|v\|_2^2\right), \quad \forall\, v \in \mathbb{R}^d.$$

Banaszczyk's Theorem by GSW algorithm

- Let $M = (v_1, \ldots, v_n)$, M-$d \times n$ matrix.
- GSW algorithm: there exists a r.v. $X \in \{-1, 1\}^n \subset \mathbb{R}^n$ s.t.

$$\mathbf{E} \exp\left(\langle MX, v \rangle\right) \leqslant \exp\left(\frac{1}{2}\|v\|_2^2\right), \quad \forall\, v \in \mathbb{R}^d.$$

- Using GSW for $v = e_i$, we get for all $i \in [d]$

$$\mathbf{P}(|\langle MX, e_i \rangle| > t) \leqslant 2 \exp\left(-\frac{1}{2} t^2\right), \quad i \in [d].$$

Banaszczyk's Theorem by GSW algorithm

- Let $M = (v_1, \ldots, v_n)$, M-$d \times n$ matrix.
- GSW algorithm: there exists a r.v. $X \in \{-1, 1\}^n \subset \mathbb{R}^n$ s.t.

$$\mathbf{E} \exp(\langle MX, v \rangle) \leqslant \exp(\frac{1}{2}\|v\|_2^2), \quad \forall\, v \in \mathbb{R}^d.$$

- Using GSW for $v = e_i$, we get for all $i \in [d]$

$$\mathbf{P}(|\langle MX, e_i \rangle| > t) \leqslant 2 \exp(-\frac{1}{2}t^2), \quad i \in [d].$$

- Hence, by the union bound

$$\mathbf{P}\left(\|MX\|_\infty > \sqrt{2}\sqrt{1 + \log d}\right) \leqslant 2d \exp(-1 - \log d) < 1.$$

Banaszczyk's Theorem by GSW algorithm

- Let $M = (v_1, \ldots, v_n)$, M-$d \times n$ matrix.
- GSW algorithm: there exists a r.v. $X \in \{-1, 1\}^n \subset \mathbb{R}^n$ s.t.

$$\mathbf{E} \exp\left(\langle MX, v\rangle\right) \leqslant \exp(\frac{1}{2}\|v\|_2^2), \quad \forall\, v \in \mathbb{R}^d.$$

- Using GSW for $v = e_i$, we get for all $i \in [d]$

$$\mathbf{P}(|\langle MX, e_i\rangle| > t) \leqslant 2\exp(-\frac{1}{2}t^2), \quad i \in [d].$$

- Hence, by the union bound

$$\mathbf{P}\left(\|MX\|_\infty > \sqrt{2}\sqrt{1 + \log d}\right) \leqslant 2d \exp\left(-1 - \log d\right) < 1.$$

- Therefore $\|MX\|_\infty \leqslant \sqrt{2}\sqrt{1 + \log d}$ for some $X \in \{-1, 1\}^n$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1, 1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1, 1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1, 1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.
- Set direction $u_t \leftarrow \mathrm{argmin}_{u \in U} \|Mu\|_2$, where $U = \{u \in \mathbb{R}^n : \; u(p_t) = 1, \; u(i) = 0 \; \forall i \notin \mathcal{A}_t\}$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.
- Set direction $u_t \leftarrow \mathrm{argmin}_{u \in U} \|Mu\|_2$, where $U = \{u \in \mathbb{R}^n : \ u(p_t) = 1, \ u(i) = 0 \ \forall i \notin \mathcal{A}_t\}$.
- $\delta^+ \leftarrow |\max \Delta|$, $\delta^- \leftarrow |\min \Delta|$, where $\Delta = \{\delta \in \mathbb{R} : \ X_t + \delta u_t \in [-1.1]^n\}$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.
- Set direction $u_t \leftarrow \mathrm{argmin}_{u \in U} \|Mu\|_2$, where $U = \{u \in \mathbb{R}^n : \; u(p_t) = 1, \; u(i) = 0 \; \forall i \notin \mathcal{A}_t\}$.
- $\delta^+ \leftarrow |\max \Delta|$, $\delta^- \leftarrow |\min \Delta|$, where $\Delta = \{\delta \in \mathbb{R} : \; X_t + \delta u_t \in [-1.1]^n\}$.
- Let

$$\delta_t \leftarrow \begin{cases} \delta^+ & \text{with you} \quad \delta^- / \left(\delta^- + \delta^+\right) \\ \delta^- & \text{with you} \quad \delta^+ / \left(\delta^- + \delta^+\right) \end{cases} .$$

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.
- Set direction $u_t \leftarrow \text{argmin}_{u \in U} \|Mu\|_2$, where $U = \{u \in \mathbb{R}^n : \ u(p_t) = 1, \ u(i) = 0 \ \forall i \notin \mathcal{A}_t\}$.
- $\delta^+ \leftarrow |\max \Delta|$, $\delta^- \leftarrow |\min \Delta|$, where $\Delta = \{\delta \in \mathbb{R} : \ X_t + \delta u_t \in [-1.1]^n\}$.
- Let

$$\boldsymbol{\delta_t} \leftarrow \begin{cases} \delta^+ & \text{with you} \quad \delta^- / \left(\delta^- + \delta^+\right) \\ \delta^- & \text{with you} \quad \delta^+ / \left(\delta^- + \delta^+\right) \end{cases} .$$

- $X_{t+1} \leftarrow X_t + \delta_t u_t$, $\mathcal{A}_{t+1} \leftarrow \{i \in [n] : |z_t(i)| < 1\}$, $t \leftarrow t+1$.

Algorithm GSW - Bansal et al. 2019

- $X_1 \in [-1,1]^n$, $t \leftarrow 1$, $\mathcal{A}_1 \leftarrow [n]$. We set the pivot: $p_0 \leftarrow n$.
- Until $\mathcal{A}_t \neq \emptyset$ do:
  1. If $p_{t-1} \notin \mathcal{A}_t$ then we set pivot $p_t = \max\{\mathcal{A}_t\}$.
  2. If $p_{t-1} \in \mathcal{A}_t$ then $p_t \leftarrow p_{t-1}$.
- Set direction $u_t \leftarrow \text{argmin}_{u \in U} \|Mu\|_2$, where $U = \{u \in \mathbb{R}^n : \ u(p_t) = 1, \ u(i) = 0 \ \forall i \notin \mathcal{A}_t\}$.
- $\delta^+ \leftarrow |\max \Delta|$, $\delta^- \leftarrow |\min \Delta|$, where $\Delta = \{\delta \in \mathbb{R} : \ X_t + \delta u_t \in [-1.1]^n\}$.
- Let

$$\delta_t \leftarrow \begin{cases} \delta^+ & \text{with you} \quad \delta^- / \left(\delta^- + \delta^+\right) \\ \delta^- & \text{with you} \quad \delta^+ / \left(\delta^- + \delta^+\right) \end{cases} .$$

- $X_{t+1} \leftarrow X_t + \delta_t u_t$, $\mathcal{A}_{t+1} \leftarrow \{i \in [n] : |z_t(i)| < 1\}$, $t \leftarrow t+1$.
- Return $X \leftarrow X_{T+1}$, where $\mathcal{A}_{T+1} = \emptyset$.

Orthogonalization

- Define the order $\sigma(1), \sigma(2), \ldots, \sigma(n)$: $\sigma(n) = p_1 = n$, then e.g.

$$\sigma(n-1) = \begin{cases} j_1 < n & \text{if} \quad j_1 - \text{frozen in step } 1 \\ n-1 & \text{if} \quad \text{in step } 1 \text{ frozen only } n \end{cases}$$

  Then, e.g. $\sigma(n-2) = j_2$ if j_2 also frozen in step 1.

Orthogonalization

- Define the order $\sigma(1), \sigma(2), \ldots, \sigma(n)$: $\sigma(n) = p_1 = n$, then e.g.

$$\sigma(n-1) = \begin{cases} j_1 < n & \text{if} \quad j_1 - \text{frozen in step } 1 \\ n-1 & \text{if} \quad \text{in step } 1 \text{ frozen only } n \end{cases}$$

  Then, e.g. $\sigma(n-2) = j_2$ if j_2 also frozen in step 1.
- Define orthogonal $w_{\sigma(1)}, w_{\sigma(2)}, \ldots w_{\sigma(n)}$

$$w_{\sigma(1)} = \frac{v_{\sigma(1)}}{\|v_{\sigma(1)}\|_2}, \quad w_{\sigma(r)} = \frac{v_{\sigma(r)} - A_r v_{\sigma(r)}}{\|v_{\sigma(r)} - A_r v_{\sigma(r)}\|_2},$$

  where $A_r = \sum_{s<r} w_{\sigma(s)} w_{\sigma(s)}^T$.

Orthogonalization

- Define the order $\sigma(1), \sigma(2), \ldots, \sigma(n)$: $\sigma(n) = p_1 = n$, then e.g.

$$\sigma(n-1) = \begin{cases} j_1 < n & \text{if} \quad j_1 - \text{frozen in step } 1 \\ n-1 & \text{if} \quad \text{in step } 1 \text{ frozen only } n \end{cases}$$

  Then, e.g. $\sigma(n-2) = j_2$ if j_2 also frozen in step 1.
- Define orthogonal $w_{\sigma(1)}, w_{\sigma(2)}, \ldots w_{\sigma(n)}$

$$w_{\sigma(1)} = \frac{v_{\sigma(1)}}{\|v_{\sigma(1)}\|_2}, \quad w_{\sigma(r)} = \frac{v_{\sigma(r)} - A_r v_{\sigma(r)}}{\|v_{\sigma(r)} - A_r v_{\sigma(r)}\|_2},$$

  where $A_r = \sum_{s<r} w_{\sigma(s)} w_{\sigma(s)}^T$.
- Note that A_r projects on $\mathrm{Lin}(v_{\sigma(1)}, \ldots, v_{\sigma(r-1)})$ and

$$Mu_t = \sum_{r=1}^{n} \langle w_{\sigma(r)}, v_{p_t} - A_t v_{p_t} \rangle w_{\sigma(r)}.$$

Orthogonalization

- Recall that $A_t = \sum_{s<t} w_{\sigma(s)} w_{\sigma(s)}^T$.

Orthogonalization

- Recall that $A_t = \sum_{s<t} w_{\sigma(s)} w_{\sigma(s)}^T$.
- Let $\ell_t = |\mathcal{A}_t|$, note that $A_t w_{\sigma(r)} = w_{\sigma(r)}$, for $r < \ell_t$.

Orthogonalization

- Recall that $A_t = \sum_{s<t} w_{\sigma(s)} w_{\sigma(s)}^T$.
- Let $\ell_t = |\mathcal{A}_t|$, note that $A_t w_{\sigma(r)} = w_{\sigma(r)}$, for $r < \ell_t$.
- For any $r \geqslant \ell_t$, the by construction $A_t w_{\sigma(r)} = 0$.

Orthogonalization

- Recall that $A_t = \sum_{s<t} w_{\sigma(s)} w_{\sigma(s)}^T$.
- Let $\ell_t = |\mathcal{A}_t|$, note that $A_t w_{\sigma(r)} = w_{\sigma(r)}$, for $r < \ell_t$.
- For any $r \geqslant \ell_t$, the by construction $A_t w_{\sigma(r)} = 0$.
- Let $g_t = \sigma^{-1}(p_t)$, then for $r > g_t$ by construction $w_{\sigma(r)} \perp v_{p_t}$.

Orthogonalization

- Recall that $A_t = \sum_{s<t} w_{\sigma(s)} w_{\sigma(s)}^T$.
- Let $\ell_t = |\mathcal{A}_t|$, note that $A_t w_{\sigma(r)} = w_{\sigma(r)}$, for $r < \ell_t$.
- For any $r \geqslant \ell_t$, the by construction $A_t w_{\sigma(r)} = 0$.
- Let $g_t = \sigma^{-1}(p_t)$, then for $r > g_t$ by construction $w_{\sigma(r)} \perp v_{p_t}$.
- Hence

$$
\begin{aligned}
Mu_t &= \sum_{r=1}^{n} \langle w_{\sigma(r)}, v_{p_t} - A_t v_{p_t} \rangle w_{\sigma(r)} \\
&= \sum_{r=1}^{n} \left[\langle w_{\sigma(r)}, v_{p_t} \rangle - \langle A_t w_{\sigma(r)}, v_{p_t} \rangle \right] w_{\sigma(r)} \\
&= \sum_{r=\ell_t}^{g_t} \langle w_{\sigma(r)}, v_{p_t} \rangle w_{\sigma(r)}.
\end{aligned}
$$

Projection matrix

- Split $V = \mathrm{Lin}(v_1, \ldots, v_n)$ into $V_1, \ldots, V_n$, where V_t is generated in step ***t***, i.e. $V_t = \emptyset$ or

$$\dim(V_t) = k+1, \quad r = \sigma^{-1}(p_t), \quad V_t = \mathrm{Lin}(w_{\sigma(r-k)}, \ldots, w_{\sigma(r)}).$$

Projection matrix

- Split $V = \mathrm{Lin}(v_1, \ldots, v_n)$ into $V_1, \ldots, V_n$, where V_t is generated in step $\boldsymbol{t}$, i.e. $V_t = \emptyset$ or

$$\dim(V_t) = k+1, \quad r = \sigma^{-1}(p_t), \quad V_t = \mathrm{Lin}(w_{\sigma(r-k)}, \ldots, w_{\sigma(r)}).$$

- We construct random projection matrix P on V.

Projection matrix

- Split $V = \mathrm{Lin}(v_1, \dots, v_n)$ into $V_1, \dots, V_n$, where V_t is generated in step $\boldsymbol{t}$, i.e. $V_t = \emptyset$ or

$$\dim(V_t) = k+1, \quad r = \sigma^{-1}(p_t), \quad V_t = \mathrm{Lin}(w_{\sigma(r-k)}, \dots, w_{\sigma(r)}).$$

- We construct random projection matrix P on V.
- Decomposition of P w.r.t. orthogonal subspaces $V_1, \dots, V_n$

$$P = \sum_{t=1}^{n} P_t, \quad P_t = 0 \text{ or } P_t = \sum_{s=0}^{k} w_{\sigma(r-s)} w_{\sigma(r-s)}^T.$$

Fixed pivot analysis

- Let p be chosen as pivot in step t_p.

Fixed pivot analysis

- Let $\boldsymbol{p}$ be chosen as pivot in step $\boldsymbol{t_p}$.
- At most $k+1$, $k \leqslant n - t_p$ vectors assigned to $\boldsymbol{p}$, i.e.:

$$w_{\sigma(r-k)}, w_{\sigma(r-k-1)}, \ldots, w_{\sigma(r)}, \quad r = \sigma^{-1}(p).$$

Fixed pivot analysis

- Let $\boldsymbol{p}$ be chosen as pivot in step $\boldsymbol{t_p}$.
- At most $k+1$, $k \leqslant n - t_p$ vectors assigned to $\boldsymbol{p}$, i.e.:

$$w_{\sigma(r-k)}, w_{\sigma(r-k-1)}, \dots, w_{\sigma(r)}, \quad r = \sigma^{-1}(p).$$

- Split set $r, r-1, \dots r-k$ into: $Q^p_{t_p-1}, Q^p_{t_p}, \dots, Q^p_n$, where $Q^p_{t_p-1} = \sigma^{-1}(p) = r$ and $Q^p_t = \mathcal{A}_t \backslash \mathcal{A}_{t+1}$, $t \geqslant t_p$, $\boldsymbol{p_t} = \boldsymbol{p}$.

Fixed pivot analysis

- Let $\boldsymbol{p}$ be chosen as pivot in step $\boldsymbol{t_p}$.
- At most $k+1$, $k \leqslant n - t_p$ vectors assigned to $\boldsymbol{p}$, i.e.:

$$w_{\sigma(r-k)}, w_{\sigma(r-k-1)}, \ldots, w_{\sigma(r)}, \quad r = \sigma^{-1}(p).$$

- Split set $r, r-1, \ldots r-k$ into: $Q^p_{t_p-1}, Q^p_{t_p}, \ldots, Q^p_n$, where $Q^p_{t_p-1} = \sigma^{-1}(p) = r$ and $Q^p_t = \mathcal{A}_t \backslash \mathcal{A}_{t+1}$, $t \geqslant t_p$, $\boldsymbol{p_t} = \boldsymbol{p}$.
- Denoting also $Q^p_t = \emptyset$ for $t \geqslant t_p$, $\boldsymbol{p_t} \neq \boldsymbol{p}$, we have for $t \geqslant t_p$

$$Mu_t = \sum_{s=t_p-1}^{t-1} \sum_{r \in Q^p_s} \langle w_{\sigma(r)}, v_p \rangle w_{\sigma(r)}.$$

Fixed pivot analysis

- Let p_t pivot in step t.

Fixed pivot analysis

- Let $\boldsymbol{p_t}$ pivot in step $\boldsymbol{t}$.
- Let $S_p = \{t \geqslant 1 : \ \ p_t = p\}$.

Fixed pivot analysis

- Let $\boldsymbol{p_t}$ pivot in step $\boldsymbol{t}$.
- Let $S_p = \{t \geqslant 1 : \; p_t = p\}$.
- Recall that, for $t \geqslant t_p$

$$Mu_t = \sum_{s=t_p-1}^{t-1} \sum_{r \in Q_s^p} \langle w_{\sigma(r)}, v_p \rangle w_{\sigma(r)}.$$

Fixed pivot analysis

- Let $\boldsymbol{p_t}$ pivot in step $\boldsymbol{t}$.
- Let $S_p = \{t \geqslant 1 : \; p_t = p\}$.
- Recall that, for $t \geqslant t_p$

$$Mu_t = \sum_{s=t_p-1}^{t-1} \sum_{r \in Q_s^p} \langle w_{\sigma(r)}, v_p \rangle w_{\sigma(r)}.$$

- Hence

$$\sum_{t \in S_p} \langle M(X_{t+1} - X_t), v \rangle = \sum_{t \in S_p} \delta_t \langle Mu_t, v \rangle$$

$$= \sum_{t=t_p}^{n} \delta_t \sum_{s=t_p-1}^{t-1} \sum_{r \in Q_s^p} \langle w_{\sigma(r)}, v_p \rangle \langle w_{\sigma(r)}, v \rangle.$$

Spielman's theorem

- Let Δ_p be generated by δ_t, $t < \min(S_p)$.

Spielman's theorem

- Let Δ_p be generated by δ_t, $t < \min(S_p)$.
- Main Observation:

Spielman's theorem

- Let Δ_p be generated by δ_t, $t < \min(S_p)$.
- Main Observation:

Theorem

Spielman et al. 2022 For each $v \in \mathbb{R}^d$

$$\mathbf{E}\left(\exp\left(\sum_{t \in S_p} \delta_t \langle M u_t, v\rangle - \frac{1}{2}\|P_p v_p\|_2^2 \cdot \|P_p v\|_2^2\right) \Big| \Delta_p\right) \leqslant 1.$$

Spielman's theorem

- Let Δ_p be generated by δ_t, $t < \min(S_p)$.
- Main Observation:

Theorem

Spielman et al. 2022 For each $v \in \mathbb{R}^d$

$$\mathbf{E}\left(\exp\left(\sum_{t\in S_p}\delta_t\langle Mu_t, v\rangle - \frac{1}{2}\|P_p v_p\|_2^2 \cdot \|P_p v\|_2^2\right)\Big|\Delta_p\right) \leqslant 1.$$

- By induction $\mathbf{E}\exp(\langle MX, v\rangle) \leqslant \exp(\frac{1}{2}\|v\|_2^2)$.

Improvement

- Recall Δ_p is generated by δ_t, $t < \min(S_p)$.

Improvement

- Recall Δ_p is generated by δ_t, $t < \min(S_p)$.
- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \;\; \beta_r = \langle w_\sigma(r), v \rangle.$$

Improvement

- Recall Δ_p is generated by δ_t, $t < \min(S_p)$.
- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \;\; \beta_r = \langle w_\sigma(r), v \rangle.$$
- We define compensator
$$F = \frac{1}{2} \left(\sum_{t=t_p-1}^{n} | \sum_{r \in Q_t^p} \alpha_r \beta_r | \right)^2 .$$

Improvement

- Recall Δ_p is generated by δ_t, $t < \min(S_p)$.
- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \;\; \beta_r = \langle w_\sigma(r), v \rangle.$$
- We define compensator
$$F = \frac{1}{2} \left(\sum_{t=t_p-1}^{n} | \sum_{r \in Q_t^p} \alpha_r \beta_r | \right)^2 .$$
- It holds true

Improvement

- Recall Δ_p is generated by δ_t, $t < \min(S_p)$.
- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \;\; \beta_r = \langle w_\sigma(r), v \rangle.$$
- We define compensator
$$F = \frac{1}{2} \left(\sum_{t=t_p-1}^{n} | \sum_{r \in Q_t^p} \alpha_r \beta_r | \right)^2 .$$
- It holds true

Lemma

B. and Godlewski 2023 For each $v \in \mathbb{R}^d$

$$\mathbf{E} \left(\exp \left(\sum_{t \in S_p} \delta_t \langle M u_t, v \rangle - F \right) \Big| \Delta_p \right) \leqslant 1.$$

Bellman's method

- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \quad \beta_r = \langle w_{\sigma(r)} v \rangle.$$

Bellman's method

- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \;\; \beta_r = \langle w_{\sigma(r)} v \rangle.$$
- We define the compensator $F = F_{n+1}$ and for $t_p \leqslant R \leqslant n$
$$F_R = \frac{1}{2} \left(\sum_{t=t_p-1}^{R-1} | \sum_{r \in Q_t^p} \alpha_r \beta_r | \right)^2 .$$

Bellman's method

- Let
$$\alpha_r = \langle w_{\sigma(r)}, v_p \rangle, \quad \beta_r = \langle w_{\sigma(r)} v \rangle.$$
- We define the compensator $F = F_{n+1}$ and for $t_p \leqslant R \leqslant n$
$$F_R = \frac{1}{2} \left(\sum_{t=t_p-1}^{R-1} | \sum_{r \in Q_t^p} \alpha_r \beta_r | \right)^2 .$$
- Idea: $g(n) \leqslant g(n-1) \leqslant \ldots g(t_p) \leqslant 1$, where
$$g(R) = \mathbf{E} \left(\exp \left(\sum_{t=t_p}^{n} \delta_t \sum_{s=t_p-1}^{(t \wedge R)-1} \sum_{r \in Q_s^p} \alpha_r \beta_r - F_R \right) \big| \Delta_p \right) .$$

Key Lemma

- Note that

Key Lemma

- Note that

Lemma

For all $a, b \in \mathbb{R}$ $\delta_R \in \{1 \pm x\}$, $x \in [-1.1]$, $\mathbf{E}\delta_R = 0$ *it holds true*

$$\mathbf{E}\exp(\delta_R(a+b)) \leqslant \exp(|a||b| + \frac{1}{2}|b|^2)\mathbf{E}\exp\left(\delta_R a\right).$$

Key Lemma

- Note that

Lemma

For all $a, b \in \mathbb{R}$ $\delta_R \in \{1 \pm x\}$, $x \in [-1.1]$, $\mathbf{E}\delta_R = 0$ *it holds true*

$$\mathbf{E}\exp(\delta_R(a+b)) \leqslant \exp(|a||b| + \frac{1}{2}|b|^2)\mathbf{E}\exp\left(\delta_R a\right).$$

- Indeed, let

$$L = \frac{(1-x)e^{-(1+x)(a+b)} + (1+x)e^{(1-x)(a+b)}}{(1-x)e^{-(1+x)a} + (1+x)e^{(1-x)a}}$$
$$= pe^{-b(1+x)} + qe^{b(1-x)} \leqslant \exp\left(\frac{1}{2}b^2 - p(1+x)b + q(1-x)b\right).$$

Key Lemma

- Note that

Lemma

For all $a, b \in \mathbb{R}$ $\delta_R \in \{1 \pm x\}$, $x \in [-1.1]$, $\mathbf{E}\delta_R = 0$ *it holds true*

$$\mathbf{E}\exp(\delta_R(a+b)) \leqslant \exp(|a||b| + \frac{1}{2}|b|^2)\mathbf{E}\exp\left(\delta_R a\right).$$

- Indeed, let

$$L = \frac{(1-x)e^{-(1+x)(a+b)} + (1+x)e^{(1-x)(a+b)}}{(1-x)e^{-(1+x)a} + (1+x)e^{(1-x)a}}$$
$$= pe^{-b(1+x)} + qe^{b(1-x)} \leqslant \exp\left(\frac{1}{2}b^2 - p(1+x)b + q(1-x)b\right).$$

- Note that

$$-p(1+x)b + q(1-x)b = \frac{(1-x^2)\sinh a}{\cosh a + x\sinh a}b \leqslant |a||b|.$$

Estimating Laplace Transform

- By induction we get

$$\mathbf{E}\exp\left(\langle MX, v\rangle - \frac{1}{2}Z_v\right) \leqslant 1,$$

where

$$Z_v = \sum_{p=1}^{n}\left(\sum_{t=t_p-1}^{n} \left|\sum_{r\in Q_t^p}\langle w_{\sigma(r)}, v_p\rangle\langle w_{\sigma(r)}, v\rangle\right|\right)^2 \leqslant \|v\|_2^2.$$

Estimating Laplace Transform

- By induction we get

$$\mathbf{E}\exp\left(\langle MX, v\rangle - \frac{1}{2}Z_v\right) \leqslant 1,$$

where

$$Z_v = \sum_{p=1}^{n}\left(\sum_{t=t_p-1}^{n}\left|\sum_{r\in Q_t^p}\langle w_{\sigma(r)}, v_p\rangle\langle w_{\sigma(r)}, v\rangle\right|\right)^2 \leqslant \|v\|_2^2.$$

- Recall $V = \mathrm{Lin}(v_1, v_2, \ldots, v_n)$. Let us define

$$m = \mathbf{E}\left|\left\{(t,p):\ \mathrm{Lin}\left(w_{\sigma(r)}:\ r\in Q_t^p\right) \neq \emptyset\right\}\right|.$$

Estimating Laplace Transform

- By induction we get

$$\mathbf{E}\exp\left(\langle MX, v\rangle - \frac{1}{2}Z_v\right) \leqslant 1,$$

where

$$Z_v = \sum_{p=1}^{n}\left(\sum_{t=t_p-1}^{n} \Big|\sum_{r\in Q_t^p}\langle w_{\sigma(r)}, v_p\rangle\langle w_{\sigma(r)}, v\rangle\Big|\right)^2 \leqslant \|v\|_2^2.$$

- Recall $V = \mathrm{Lin}(v_1, v_2, \dots, v_n)$. Let us define

$$m = \mathbf{E}\Big|\left\{(t,p):\ \mathrm{Lin}\left(w_{\sigma(r)}:\ r\in Q_t^p\right)\neq\emptyset\right\}\Big|.$$

- Note that $m \leqslant \min\{\mathbf{E}T, \dim V\}$.

Estimating Laplace Transform

- By induction we get

$$\mathbf{E} \exp\left(\langle MX, v\rangle - \frac{1}{2} Z_v \right) \leqslant 1,$$

where

$$Z_v = \sum_{p=1}^{n} \left(\sum_{t=t_p-1}^{n} \Big| \sum_{r \in Q_t^p} \langle w_{\sigma(r)}, v_p\rangle \langle w_{\sigma(r)}, v\rangle \Big| \right)^2 \leqslant \|v\|_2^2.$$

- Recall $V = \mathrm{Lin}(v_1, v_2, \ldots, v_n)$. Let us define

$$m = \mathbf{E} \Big| \left\{ (t, p) : \ \mathrm{Lin}\left(w_{\sigma(r)} : \ r \in Q_t^p \right) \neq \emptyset \right\} \Big|.$$

- Note that $m \leqslant \min\{\mathbf{E}T, \dim V\}$.
- It is easy to see that $Z_{e_i} = Z_{-e_i}$ and $\sum_{i=1}^{d} \mathbf{E} Z_{e_i} \leqslant m$.

Estimating Laplace transform

- Recall that $\mathbf{E}\langle MX, v\rangle = 0$ and

$$\mathbf{E}\exp\left(\lambda\langle MX, v\rangle - \frac{1}{2}\lambda^2 Z_v\right) \leqslant 1.$$

Estimating Laplace transform

- Recall that $\mathbf{E}\langle MX, v\rangle = 0$ and

$$\mathbf{E} \exp\left(\lambda\langle MX, v\rangle - \frac{1}{2}\lambda^2 Z_v\right) \leqslant 1.$$

- The analysis of $\lambda \to 0$ gives

$$\mathbf{E}|\langle MX, v\rangle|^2 \leqslant \frac{1}{2}\mathbf{E}Z_v.$$

Estimating Laplace transform

- Recall that $\mathbf{E}\langle MX, v\rangle = 0$ and

$$\mathbf{E}\exp\left(\lambda\langle MX, v\rangle - \frac{1}{2}\lambda^2 Z_v\right) \leqslant 1.$$

- The analysis of $\lambda \to 0$ gives

$$\mathbf{E}|\langle MX, v\rangle|^2 \leqslant \frac{1}{2}\mathbf{E}Z_v.$$

- Consider $\varphi(x) = (e^x - x - 1)\,1_{x>0}$, we get for $\lambda > 0$

$$\mathbf{E}\varphi(\lambda\langle MX, e_i\rangle - \frac{1}{2}\lambda^2) \leqslant \mathbf{E}Z_{e_i}\left[D\lambda^2\left(1 + D\lambda^2\right)\right].$$

Estimating Laplace transform

- Recall that $\mathbf{E}\langle MX, v\rangle = 0$ and

$$\mathbf{E}\exp\left(\lambda\langle MX, v\rangle - \frac{1}{2}\lambda^2 Z_v\right) \leqslant 1.$$

- The analysis of $\lambda \to 0$ gives

$$\mathbf{E}|\langle MX, v\rangle|^2 \leqslant \frac{1}{2}\mathbf{E}Z_v.$$

- Consider $\varphi(x) = (e^x - x - 1)\,1_{x>0}$, we get for $\lambda > 0$

$$\mathbf{E}\varphi(\lambda\langle MX, e_i\rangle - \frac{1}{2}\lambda^2) \leqslant \mathbf{E}Z_{e_i}\left[D\lambda^2\left(1 + D\lambda^2\right)\right].$$

- Thus for $\lambda \sim \sqrt{1 + \log m}$

$$\sum_{i=1}^{d}\mathbf{P}(|\langle MX, e_i\rangle| \geqslant C\sqrt{1 + \log m}) < \frac{1}{m}\sum_{i=1}^{d}\mathbf{E}Z_{e_i} < 1.$$

Conclusion

- Banaszczyk's theorem can be improved:

Conclusion

- Banaszczyk's theorem can be improved:

Theorem

B. and Godlewski 2023 There is $(\varepsilon_i)_{i=1}^n \in \{\pm 1\}^n$ *such that*

$$\|\sum_{i=1}^n \varepsilon_i v_i\|_\infty \leqslant C\sqrt{1 + \log m},$$

where $m \leqslant \min\{\mathbf{E}T, \dim(V)\}$, *where* $V = \mathrm{Lin}(v_1, v_2, \ldots, v_n)$.

Conclusion

- Banaszczyk's theorem can be improved:

Theorem

B. and Godlewski 2023 There is $(\varepsilon_i)_{i=1}^n \in \{\pm 1\}^n$ *such that*

$$\|\sum_{i=1}^{n} \varepsilon_i v_i\|_\infty \leqslant C\sqrt{1+\log m},$$

where $m \leqslant \min\{\mathbf{E}T, \dim(V)\}$*, where* $V = \mathrm{Lin}(v_1, v_2, \ldots, v_n)$*.*

- The analysis of moments

$$\mathbf{E}|\langle MX, v\rangle|^p?$$

Smoothed Kómlos

- Smoothing the Kómlos hypothesis

Smoothed Kómlos

- Smoothing the Kómlos hypothesis

Theorem

Meka et al. 2022 Let $\sigma > 0$, $n \sim \frac{d \log d}{\sigma^{4/3}}$*. Then with positive prob. there is* $\varepsilon = (\varepsilon_i)_{i=1}^n \in \{\pm 1\}_{i=1}^n$

$$\|(M + R)\varepsilon\|_\infty \leqslant \frac{1}{d^c},$$

where R has inputs i.i.d. from the distribution $\mathcal{N}(0, \sigma^2/d)$*.*

Smoothed Kómlos

- Smoothing the Kómlos hypothesis

Theorem

Meka et al. 2022 Let $\sigma > 0$*,* $n \sim \frac{d \log d}{\sigma^{4/3}}$*. Then with positive prob. there is* $\varepsilon = (\varepsilon_i)_{i=1}^n \in \{\pm 1\}_{i=1}^n$

$$\|(M + R)\varepsilon\|_\infty \leqslant \frac{1}{d^c},$$

where R has inputs i.i.d. from the distribution $\mathcal{N}(0, \sigma^2/d)$*.*

- We can replace in the analysis d by m.

Thanks for your attention!