The Recursive Hessian Sketch for Adaptive Filtering

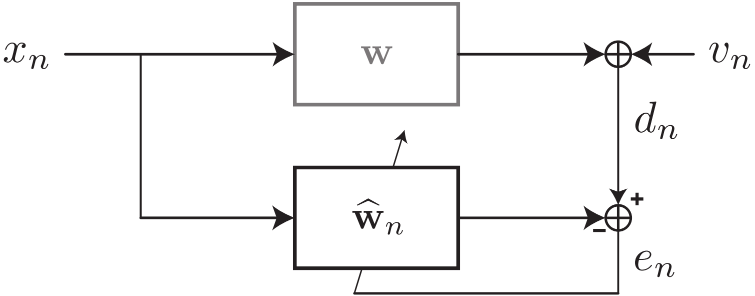
Robin Scheibler   Martin Vetterli

School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne, Switzerland

ICASSP
March 25, 2016

LCAV

Adaptive filters



- Unknown filter $\mathbf{w}$
- Access to x_n, d_n
- Estimate of unknown filter $\widehat{\mathbf{w}}_n$
- Estimation error $e_n = d_n - \widehat{\mathbf{w}}_n^\top \mathbf{x}$

1st order methods – Least mean squares (LMS)

- Stochastic gradient descent:

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \mathbb{E}|e_n|^2, \qquad \widehat{\mathbf{w}}_n = \widehat{\mathbf{w}}_{n-1} + \mu e_n \mathbf{x}_n$$

- Cheap, robust

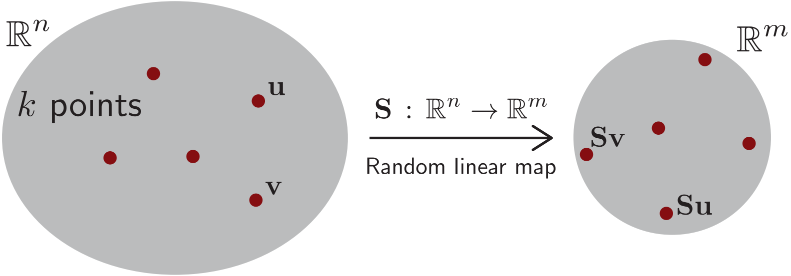
2nd order methods – Recursive Least Squares (RLS)

- Least squares problem:

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \left\| \mathbf{\Lambda}_n^{1/2} (\mathbf{X}_n \mathbf{w} - \mathbf{d}_n) \right\|^2$$

- Complex, faster convergence, lower residual

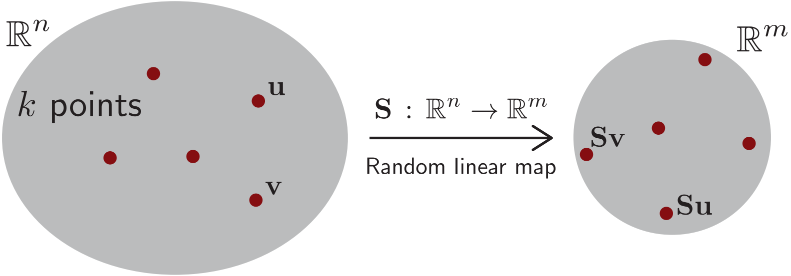
Theorem (Johnson-Lindenstrauss lemma)



Distances are preserved whp.

$$(1-\epsilon)\|\mathbf{u}-\mathbf{v}\|^2 \leq \|\mathbf{Su}-\mathbf{Sv}\|^2 \leq (1+\epsilon)\|\mathbf{u}-\mathbf{v}\|^2$$

Theorem (Johnson-Lindenstrauss lemma)

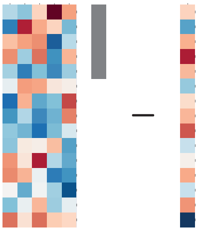


Distances are preserved whp.

$$(1-\epsilon)\|\mathbf{u}-\mathbf{v}\|^2 \leq \|\mathbf{Su}-\mathbf{Sv}\|^2 \leq (1+\epsilon)\|\mathbf{u}-\mathbf{v}\|^2$$

Application to solving least squares problem

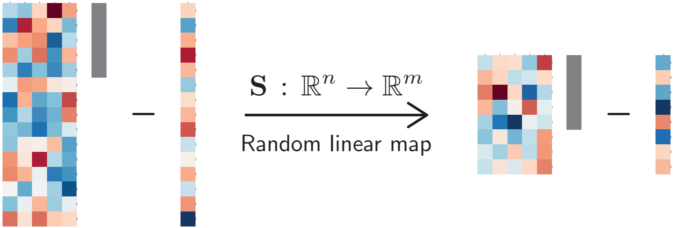
$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$$



- Smaller system to solve!
- The J-L lemma implies $\|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|^2 \leq (1+\epsilon)\|\mathbf{A}\mathbf{x}^{\mathrm{LS}} - \mathbf{b}\|^2$
- **But** no good bound on solution error

Application to solving least squares problem

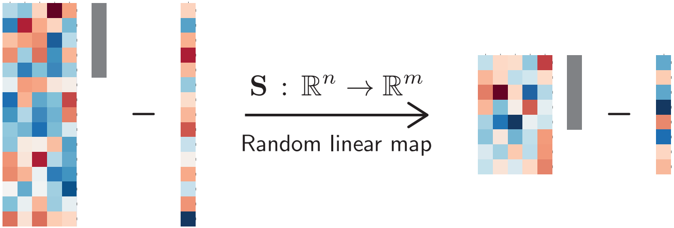
$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \qquad\qquad \|\mathbf{S}\mathbf{A}\mathbf{x} - \mathbf{S}\mathbf{b}\|^2$$



- Smaller system to solve!
- The J-L lemma implies $\|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|^2 \leq (1+\epsilon)\|\mathbf{A}\mathbf{x}^{\mathrm{LS}} - \mathbf{b}\|^2$
- **But** no good bound on solution error

Application to solving least squares problem

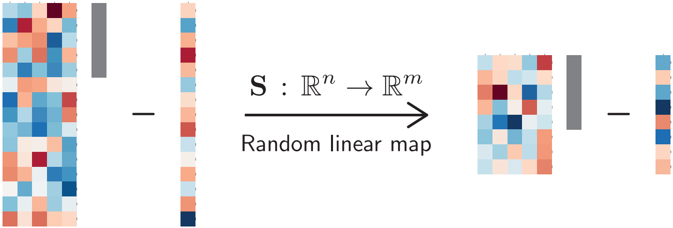
$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$ $\|\mathbf{S}\mathbf{A}\mathbf{x} - \mathbf{S}\mathbf{b}\|^2$



- Smaller system to solve!
- The J-L lemma implies $\|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|^2 \leq (1+\epsilon)\|\mathbf{A}\mathbf{x}^{\mathsf{LS}} - \mathbf{b}\|^2$
- **But** no good bound on solution error

Application to solving least squares problem

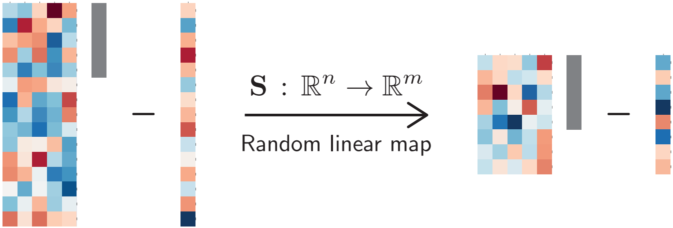
$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$ $\|\mathbf{S}\mathbf{A}\mathbf{x} - \mathbf{S}\mathbf{b}\|^2$



- Smaller system to solve!
- The J-L lemma implies $\|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|^2 \leq (1+\epsilon)\|\mathbf{A}\mathbf{x}^{\mathsf{LS}} - \mathbf{b}\|^2$
- **But** no good bound on solution error

Application to solving least squares problem

$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$ $\qquad\qquad$ $\|\mathbf{S}\mathbf{A}\mathbf{x} - \mathbf{S}\mathbf{b}\|^2$



- Smaller system to solve!
- The J-L lemma implies $\|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|^2 \leq (1+\epsilon)\|\mathbf{A}\mathbf{x}^{\mathsf{LS}} - \mathbf{b}\|^2$
- **But** no good bound on solution error

Apply sketching to the RLS algorithm.

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \left\| \mathbf{\Lambda}_n^{1/2} (\mathbf{X}_n \mathbf{w} - \mathbf{d}_n) \right\|^2$$

Wish list

- As good as RLS
- With less computations
- Good convergence

Apply sketching to the RLS algorithm.

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \left\| \mathbf{\Lambda}_n^{1/2} (\mathbf{X}_n \mathbf{w} - \mathbf{d}_n) \right\|^2$$

Wish list

- As good as RLS
- With less computations
- Good convergence

Apply sketching to the RLS algorithm.

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \left\| \boldsymbol{\Lambda}_n^{1/2} (\mathbf{X}_n \mathbf{w} - \mathbf{d}_n) \right\|^2$$

Wish list

- As good as RLS
- With less computations
- Good convergence

Apply sketching to the RLS algorithm.

$$\widehat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \left\| \mathbf{\Lambda}_n^{1/2} (\mathbf{X}_n \mathbf{w} - \mathbf{d}_n) \right\|^2$$

Wish list

- As good as RLS
- With less computations
- Good convergence

1. The iterative Hessian sketch
2. The recursive least squares
3. The recursive Hessian sketch

The iterative Hessian sketch

The Hessian sketch for least-squares

M. Pilanci, M. J. Wainwright, *Iterative Hessian sketch: Fast and accurate solution approximation for constrained least-squares*, 2014.

Goal

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2, \qquad \begin{array}{lcl} \mathbf{A} & : & \text{data matrix} \\ \mathbf{b} & : & \text{response vector} \end{array}$$

The Hessian sketch

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top \mathbf{b}\right)^\top \mathbf{x}$$

Sketch **only** data matrix, then

$$\frac{\|\mathbf{x}^{\mathsf{LS}} - \tilde{\mathbf{x}}\|_{\mathbf{A}}}{\|\mathbf{x}^{\mathsf{LS}}\|_{\mathbf{A}}} \leq \delta$$

The Hessian sketch for least-squares

M. Pilanci, M. J. Wainwright, *Iterative Hessian sketch: Fast and accurate solution approximation for constrained least-squares*, 2014.

Goal

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2,$$

$\mathbf{A}$: data matrix

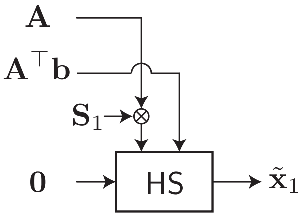
$\mathbf{b}$: response vector

The Hessian sketch

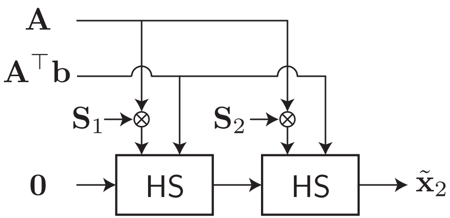
$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top \mathbf{b}\right)^\top \mathbf{x}$$

Sketch **only** data matrix, then

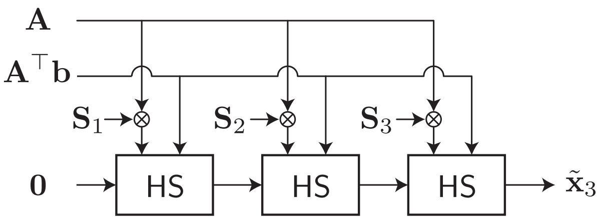
$$\frac{\|\mathbf{x}^{\mathsf{LS}} - \tilde{\mathbf{x}}\|_{\mathbf{A}}}{\|\mathbf{x}^{\mathsf{LS}}\|_{\mathbf{A}}} \leq \delta$$



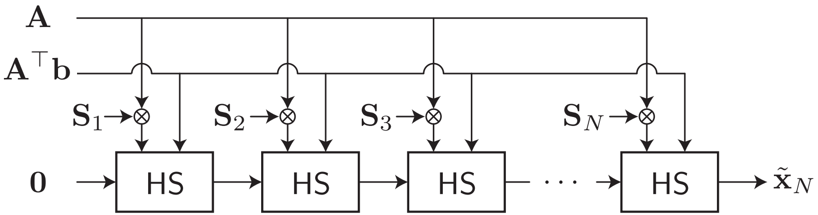
Relative error: δ



Relative error: δ^2



Relative error: δ^3



Relative error: ϵ in $N = \log(1/\epsilon)$ iterations

- Sketch data matrix, not the response vector
- ϵ-approx of LS in $\log(1/\epsilon)$ iterations
- Save computational cost of $\mathbf{A}^T\mathbf{A}$

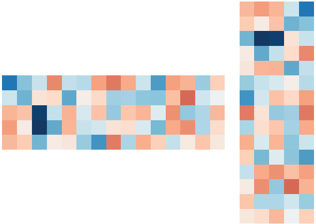
Iterative Hessian sketch : summary

- Sketch data matrix, not the response vector
- ϵ-approx of LS in $\log(1/\epsilon)$ iterations
- Save computational cost of $\mathbf{A}^T\mathbf{A}$

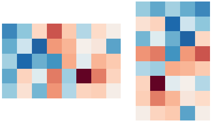
Iterative Hessian sketch : summary

- Sketch data matrix, not the response vector
- ϵ-approx of LS in $\log(1/\epsilon)$ iterations
- Save computational cost of $\mathbf{A}^T\mathbf{A}$

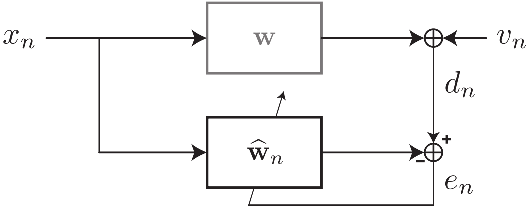
$$\mathbf{A}^\top \mathbf{A}$$



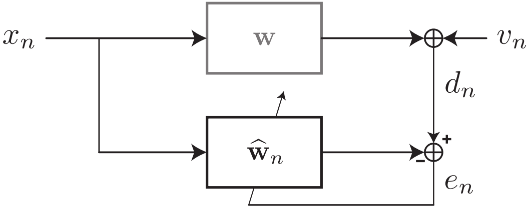
$$(\mathbf{SA})^\top \mathbf{SA}$$



The recursive least squares

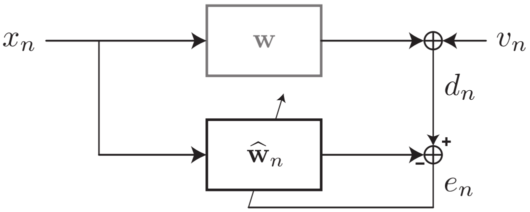

x_n
$\mathbf{w}$
v_n
d_n
$\widehat{\mathbf{w}}_n$
+
−
e_n

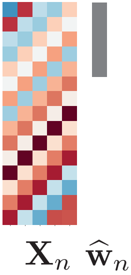
Exponentially weighted least squares

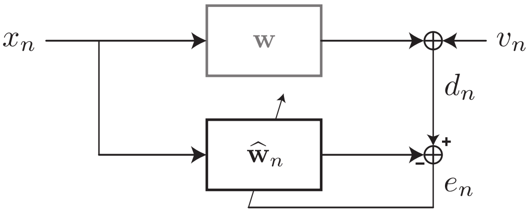


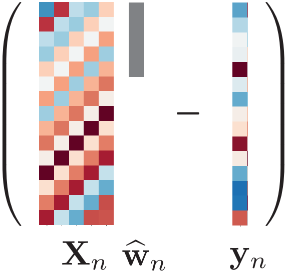
$\widehat{\mathbf{w}}_n$

Exponentially weighted least squares

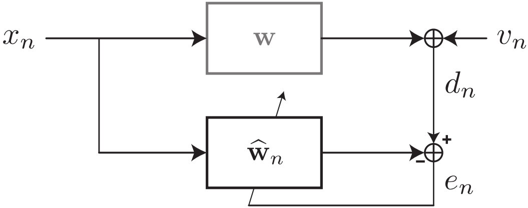


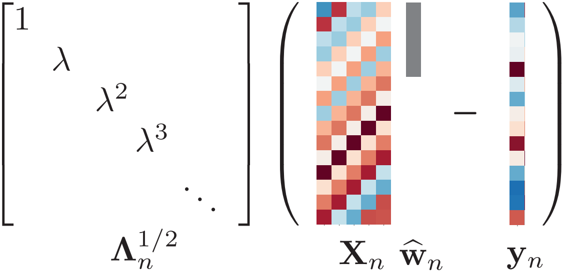



x_n
$\mathbf{w}$
v_n
d_n
$\widehat{\mathbf{w}}_n$
+
-
e_n

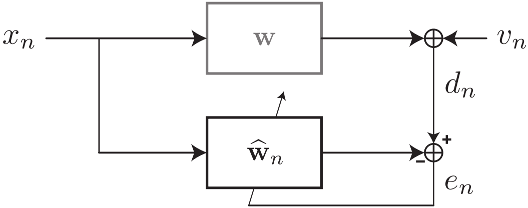

$-$
$\mathbf{X}_n$
$\widehat{\mathbf{w}}_n$
$\mathbf{y}_n$

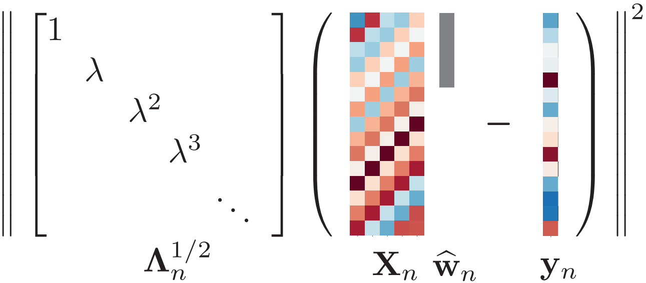
Exponentially weighted least squares





Exponentially weighted least squares





RLS filter update

$$\widehat{\mathbf{w}}_n = \underbrace{\left(\mathbf{X}_n^\top \boldsymbol{\Lambda}_n \mathbf{X}_n\right)}_{\mathbf{R}_n}^{-1} \underbrace{\mathbf{X}_n^\top \boldsymbol{\Lambda}_n \mathbf{d}_n}_{\mathbf{y}_n}$$

Data update

$$\mathbf{X}_{n+1} = \begin{bmatrix} \mathbf{x}^\top \\ \mathbf{X}_n \end{bmatrix} \quad \mathbf{d}_{n+1} = \begin{bmatrix} d \\ \mathbf{d}_n \end{bmatrix}$$

RLS filter update

$$\widehat{\mathbf{w}}_{n+1} = \underbrace{\left(\lambda \mathbf{R}_n + \mathbf{x}\mathbf{x}^\top\right)}_{\text{rank-1 update!}}^{-1} \underbrace{\left(\lambda \mathbf{y}_n + \mathbf{x} d\right)}_{\mathbf{y}_{n+1}}$$

RLS filter update

$$\widehat{\mathbf{w}}_n = \underbrace{\left(\mathbf{X}_n^\top \boldsymbol{\Lambda}_n \mathbf{X}_n\right)}_{\mathbf{R}_n}^{-1} \underbrace{\mathbf{X}_n^\top \boldsymbol{\Lambda}_n \mathbf{d}_n}_{\mathbf{y}_n}$$

Data update

$$\mathbf{X}_{n+1} = \begin{bmatrix} \mathbf{x}^\top \\ \mathbf{X}_n \end{bmatrix} \qquad \mathbf{d}_{n+1} = \begin{bmatrix} d \\ \mathbf{d}_n \end{bmatrix}$$

RLS filter update

$$\widehat{\mathbf{w}}_{n+1} = \underbrace{\left(\lambda \mathbf{R}_n + \mathbf{x}\mathbf{x}^\top\right)}_{\text{rank-1 update!}}^{-1} \underbrace{(\lambda \mathbf{y}_n + \mathbf{x} d)}_{\mathbf{y}_{n+1}}$$

RLS filter update

$$\widehat{\mathbf{w}}_n = \underbrace{\left(\mathbf{X}_n^\top \mathbf{\Lambda}_n \mathbf{X}_n\right)}_{\mathbf{R}_n}^{-1} \underbrace{\mathbf{X}_n^\top \mathbf{\Lambda}_n \mathbf{d}_n}_{\mathbf{y}_n}$$

Data update

$$\mathbf{X}_{n+1} = \begin{bmatrix} \mathbf{x}^\top \\ \mathbf{X}_n \end{bmatrix} \qquad \mathbf{d}_{n+1} = \begin{bmatrix} d \\ \mathbf{d}_n \end{bmatrix}$$

RLS filter update

$$\widehat{\mathbf{w}}_{n+1} = \underbrace{\left(\lambda \mathbf{R}_n + \mathbf{x}\mathbf{x}^\top\right)}_{\text{rank-1 update!}}^{-1} \underbrace{(\lambda \mathbf{y}_n + \mathbf{x} d)}_{\mathbf{y}_{n+1}}$$

Summary of RLS algorithm

- Solve LS at each step
- Update solution with matrix inversion lemma
- Cost quadratic in filter length

- Solve LS at each step
- Update solution with matrix inversion lemma
- Cost quadratic in filter length

Summary of RLS algorithm

- Solve LS at each step
- Update solution with matrix inversion lemma
- Cost quadratic in filter length

The recursive Hessian sketch

- Recall the Hessian sketch $(\mathbf{A} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{X}_n,\ \mathbf{b} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{d}_n)$

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

Caveat: IHS proof does not cover this sketch (yet)

- Recall the Hessian sketch ($\mathbf{A} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

Caveat: IHS proof does not cover this sketch (yet)

- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

Caveat: IHS proof does not cover this sketch (yet)

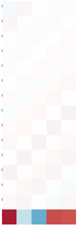
- Recall the Hessian sketch ($\mathbf{A} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

$b_1 = 1$



Caveat: IHS proof does not cover this sketch (yet)

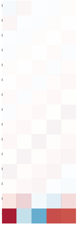
- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

$$b_2 = 0$$



Caveat: IHS proof does not cover this sketch (yet)

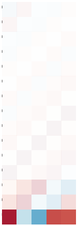
- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

$b_3 = 0$



Caveat: IHS proof does not cover this sketch (yet)

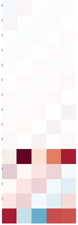
- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

$b_4 = 1$



Caveat: IHS proof does not cover this sketch (yet)

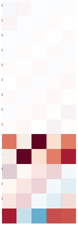
- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

$b_5 = 1$



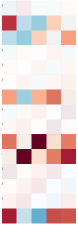
Caveat: IHS proof does not cover this sketch (yet)

- Recall the Hessian sketch ($\mathbf{A} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \boldsymbol{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$



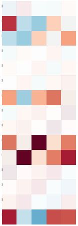
Caveat: IHS proof does not cover this sketch (yet)

- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$



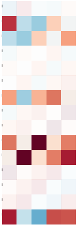
Caveat: IHS proof does not cover this sketch (yet)

- Recall the Hessian sketch ($\mathbf{A} = \mathbf{\Lambda}_n^{1/2}\mathbf{X}_n$, $\mathbf{b} = \mathbf{\Lambda}_n^{1/2}\mathbf{d}_n$)

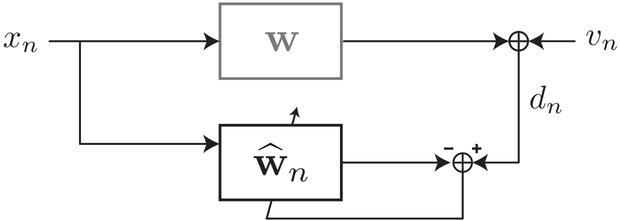
$$\tilde{\mathbf{w}}_n = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{S}_n\mathbf{A}\mathbf{x}\|^2 - \left(\mathbf{A}^\top\mathbf{b}\right)^\top \mathbf{x}$$

- Random row sampling : $\mathbf{S}_n$, fixed aspect ratio $q = \frac{m}{n}$

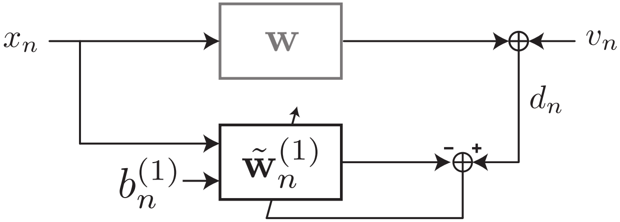
$$\mathbf{S}_n = \begin{bmatrix} b_n & 0 \\ 0 & \mathbf{S}_{n-1} \end{bmatrix}, \quad b_n = \begin{cases} 1 & \text{w.p. } q \\ 0 & \text{w.p. } 1-q \end{cases}$$

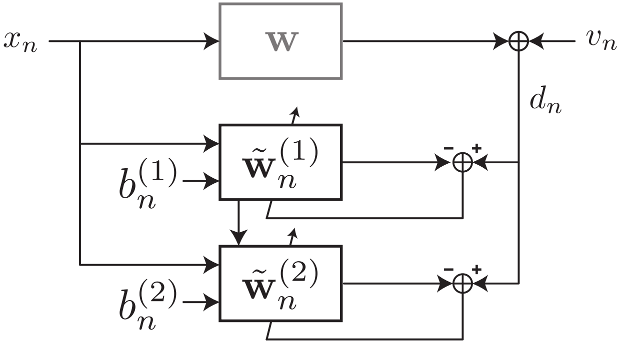


Caveat: IHS proof does not cover this sketch (yet)

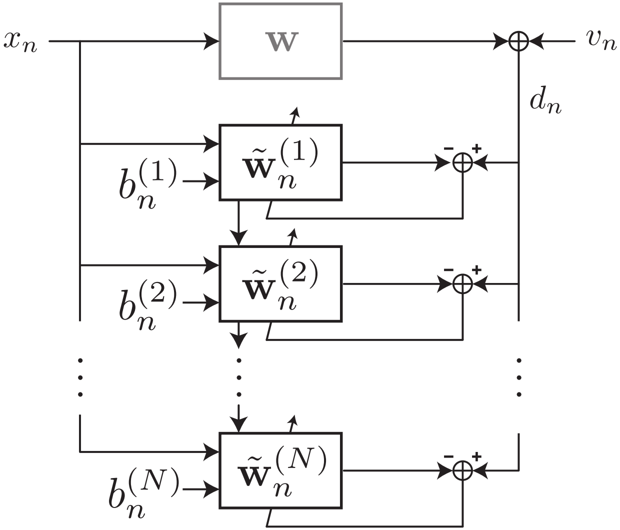

x_n
$\mathbf{w}$
v_n
d_n
$\widehat{\mathbf{w}}_n$
−
+

The recursive Hessian sketch (RHS)




x_n
$\mathbf{w}$
v_n
d_n
$\tilde{\mathbf{w}}_n^{(1)}$
$b_n^{(1)}$
$\tilde{\mathbf{w}}_n^{(2)}$
$b_n^{(2)}$
-
+
-
+

The recursive Hessian sketch (RHS)



Summary of RHS algorithm

- Apply Hessian sketch to RLS
- Update inverse matrix wp q
- Cascade N sketched RLS

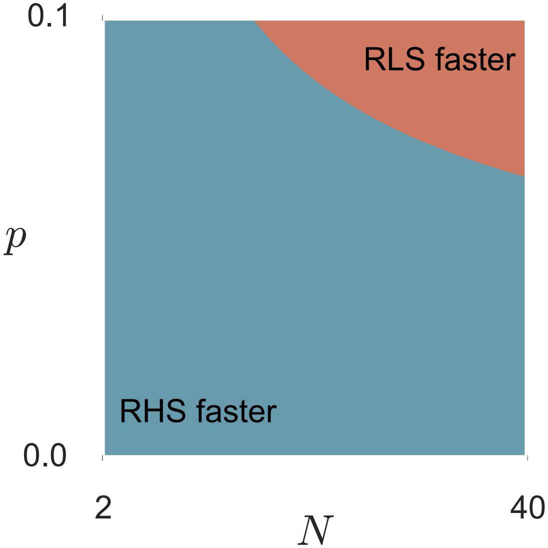
Summary of RHS algorithm

- Apply Hessian sketch to RLS
- Update inverse matrix wp q
- Cascade N sketched RLS

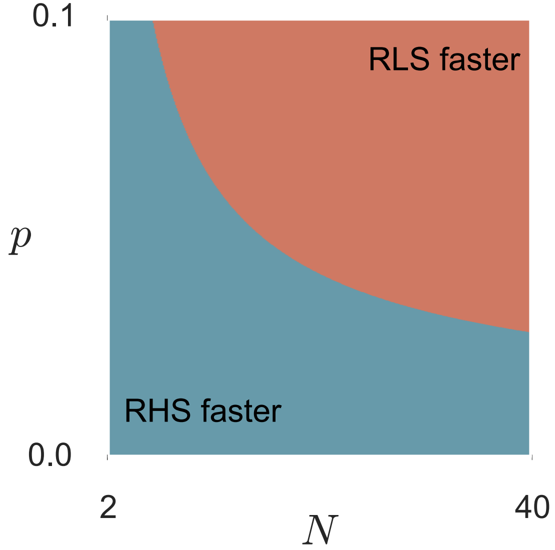
Summary of RHS algorithm

- Apply Hessian sketch to RLS
- Update inverse matrix wp q
- Cascade N sketched RLS

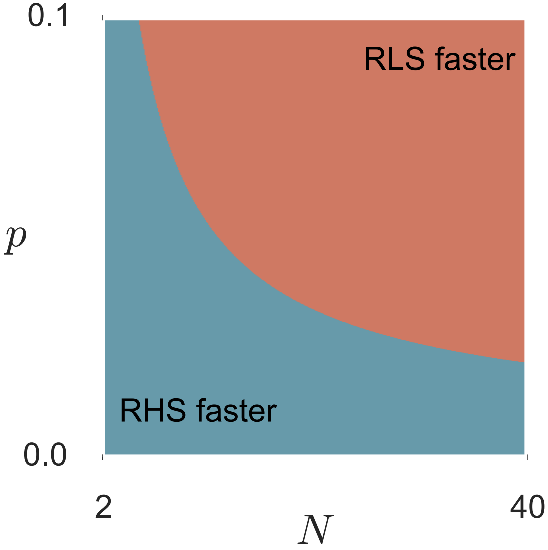
Filter length : 10



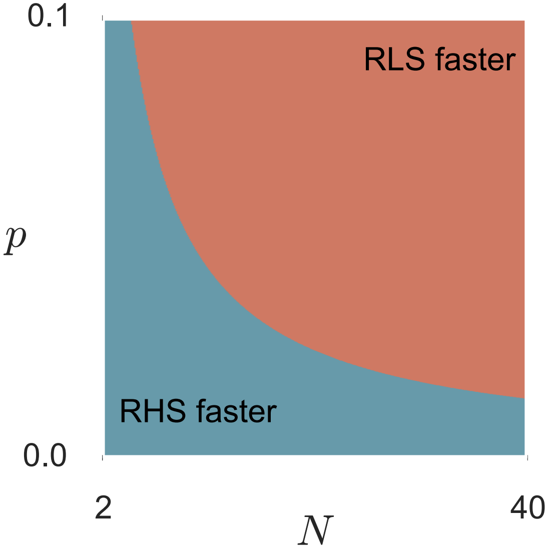
Filter length : 50



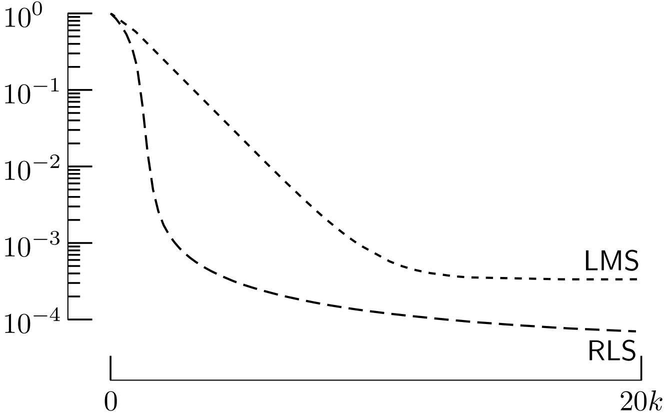
Filter length : 100



Filter length : 1000

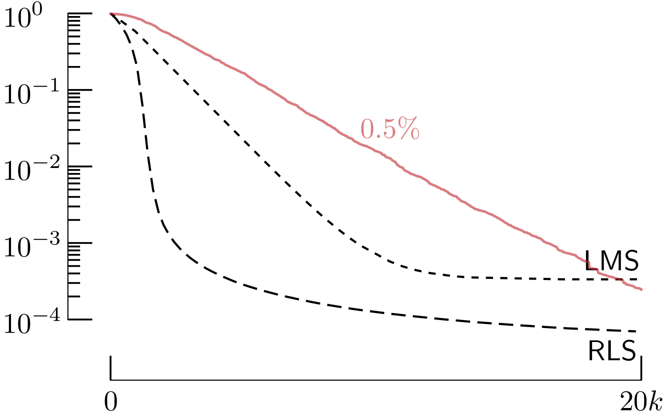


Filter length 1000, $N = 5$, 300 realizations

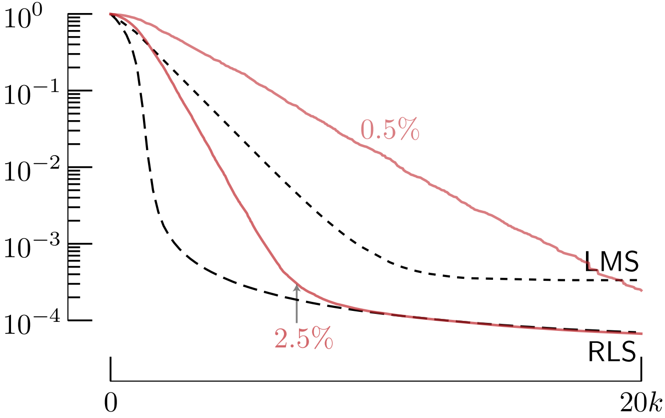


Simulation results — MSE, SNR 30dB

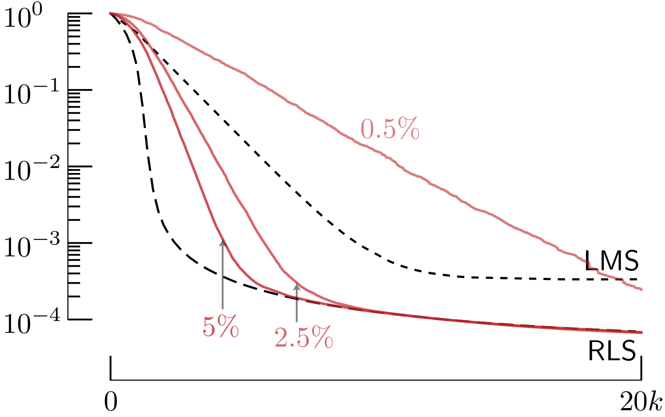
Filter length 1000, $N = 5$, 300 realizations



Filter length 1000, $N = 5$, 300 realizations

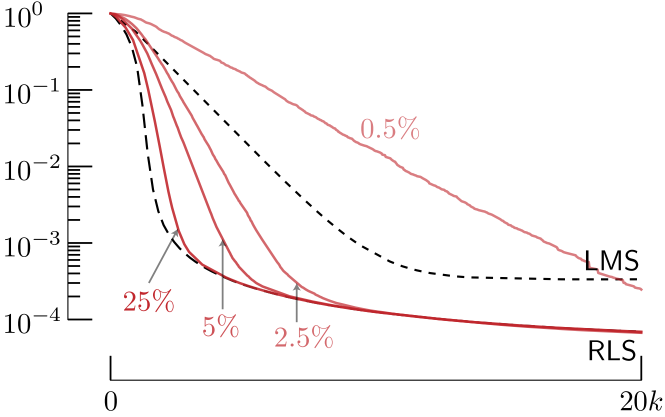


Filter length 1000, $N = 5$, 300 realizations



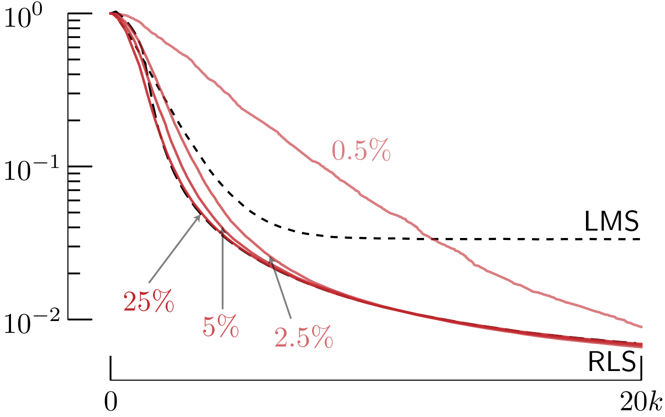
Simulation results — MSE, SNR 30dB

Filter length 1000, $N = 5$, 300 realizations



Simulation results — MSE, SNR 10dB

Filter length 1000, $N = 5$, 300 realizations



Contributions

- A sketched adaptive filter converging to RLS solution
- Lower computational complexity
- Extensive simulation

What's next ?

- Proof of IHS for random row sampling
- Experiments with non-stationary input (e.g. audio, speech)
- Investigate tracking behavior

Contributions

- A sketched adaptive filter converging to RLS solution
- Lower computational complexity
- Extensive simulation

What's next ?

- Proof of IHS for random row sampling
- Experiments with non-stationary input (e.g. audio, speech)
- Investigate tracking behavior

Thanks for your attention!

Code and figures available at

`http://github.com/LCAV/SketchRLS/`