Recursive Estimation of Room Impulse Responses with Energy Conservation Constraints

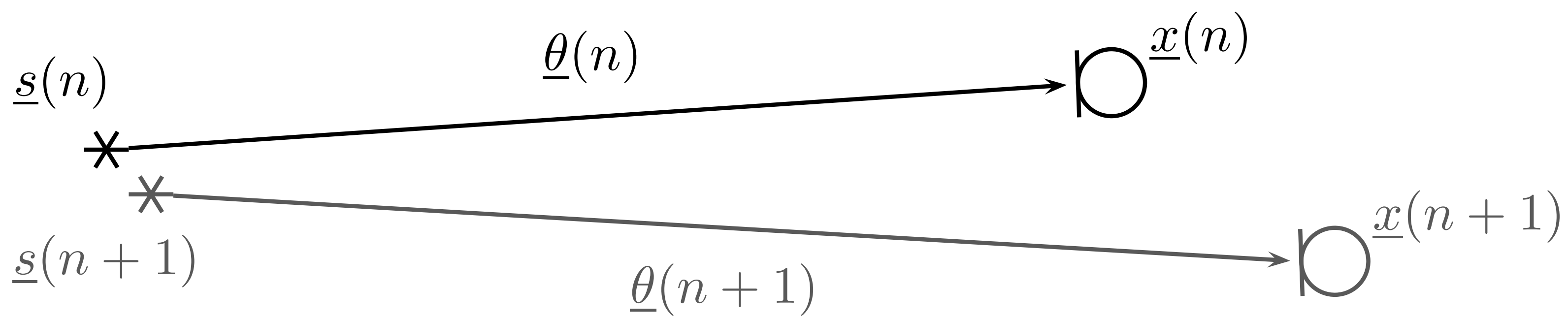
Stefan Uhlich and Bin Yang

Chair of System Theory and Signal Processing
University of Stuttgart

May 25, 2011

Introduction and Motivation (I)

- Goal: Track time-varying room impulse response (RIR) between source and microphone



- Signal model:

$$\underline{x}(n) = \underline{\theta}(n) * \underline{s}(n) + \underline{z}(n), \qquad n = 1, \ldots, N$$

with

$\underline{x}(n) \in \mathbb{R}^K$	...	measured microphone signal
$\underline{\theta}(n) \in \mathbb{R}^M$	...	room impulse response
$\underline{s}(n) \in \mathbb{R}^S$	...	known source signal
$\underline{z}(n) \in \mathbb{R}^K$	...	Gaussian noise, $\underline{z}(n) \sim \mathcal{N}(\underline{0}, \mathbf{C}(n))$, temporally uncorr.

Introduction and Motivation (II)

- Signal model (in Matrix notation):

$$\underline{x}(n) = \mathbf{\Theta}(n)\underline{s}(n) + \underline{z}(n) = \mathbf{S}(n)\underline{\theta}(n) + \underline{z}(n), \qquad n = 1, \dots, N$$

where $\mathbf{\Theta}(n) \in \mathbb{R}^{K\times S}$ and $\mathbf{S}(n) \in \mathbb{R}^{K\times M}$ are Toeplitz matrices with

$$\mathbf{\Theta}(n) = \begin{bmatrix} \underline{\theta}(n) & & & \mathbf{0} \\ & \underline{\theta}(n) & & \\ & & \ddots & \\ \mathbf{0} & & & \underline{\theta}(n) \end{bmatrix} \qquad \mathbf{S}(n) = \begin{bmatrix} \underline{s}(n) & & & \mathbf{0} \\ & \underline{s}(n) & & \\ & & \ddots & \\ \mathbf{0} & & & \underline{s}(n) \end{bmatrix}$$

- Popular solution approach: Weighted least squares (WLS) estimator

$$\hat{\underline{\theta}}_{\mathsf{WLS}}(n) = \arg\min_{\underline{\theta}} \sum_{i=0}^{n} \beta^{n-i} \left(\underline{x}(i) - \mathbf{S}(i)\underline{\theta}\right)^T \mathbf{C}(i)^{-1} \left(\underline{x}(i) - \mathbf{S}(i)\underline{\theta}\right)$$

where $0 \leq \beta \leq 1$ is the exponential forgetting factor

Introduction and Motivation (III)

- Idea: Improve estimation performance by incorporating a priori information

- Open Questions:
  - **What a priori information ($=$ constraint) can we exploit?**
    - Energy conservation constraint $\|\mathbf{\Theta}(n)\underline{s}(n)\|^2 \leq \|\underline{s}(n)\|^2$
  - **How to efficiently include the constraint into tracking algorithm?**
    - Recursive constrained maximum likelihood (RCML) estimator
    - Recursive affine minimax (RAMX) estimator
    - Recursive minimum mean squared error (RMMSE) estimator

Contents

Energy Conservation Constraint (I)

- Expressing the energy conservation mathematically, we need to ensure that

$$\|\mathbf{\Theta}(n)\underline{s}(n)\|^2 = \underline{s}(n)^T\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)\underline{s}(n) \ \leq \ \underline{s}(n)^T\underline{s}(n) = \|\underline{s}(n)\|^2$$

  i.e.

  "signal energy at microphone" $\leq$ "signal energy at source"

  which implies

$$\mathbf{\Theta}(n)^T\mathbf{\Theta}(n) - \mathbf{I} \preceq \mathbf{0}$$

- Condition has to hold for all signal lengths S
  - if condition fulfilled for particular S_0 then also fulfilled for all $S < S_0$
  - therefore, it is sufficient to consider case $S \to \infty$

Energy Conservation Constraint (II)

- Two equivalent representations of the set $\Theta = \{\underline{\theta} : \mathbf{\Theta}(n)^T \mathbf{\Theta}(n) - \mathbf{I} \preceq \mathbf{0}\}$:
  - LMI representation: Using Schur's lemma, the constraint is

$$\begin{bmatrix} \mathbf{I} & \mathbf{\Theta}(n)^T \\ \mathbf{\Theta}(n) & \mathbf{I} \end{bmatrix} \succeq \mathbf{0}$$

    i.e. Θ is a convex set.

  - Frequency domain representation:
    - Use equivalence of the eigenvalues of $\mathbf{\Theta}(n)^T \mathbf{\Theta}(n)$ (bandlimited Toeplitz matrix) and corresponding circulant matrix for $S \to \infty$
    - Constraint can be written as

$$|\mathit{\Theta}(\omega, n)|^2 = \left| \sum_{m=0}^{M-1} \theta_m(n) e^{-j\omega m} \right|^2 \leq 1 \; \forall \, \omega \in [0, 2\pi)$$

      where $\mathit{\Theta}(\omega, n)$ is the room frequency response

Energy Conservation Constraint (III)

- All three recursive estimators are based on frequency domain representation

- Approximation of the frequency domain representation:

  - Use DFT to evaluate $\Theta(\omega, n)$ at discrete frequencies ω_l

$$\omega_l = \frac{2\pi l}{L} \quad \text{with} \quad l = 0, \ldots, L-1$$

  where $L \geq M$.

  - Let $\mathbf{P} \in \mathbb{R}^{M \times L}$ be matrix consisting of the first M rows of $\mathbf{I}$. Then,

$$|\Theta(\omega_l, n)|^2 = \underline{\theta}(n)^T \mathbf{P} \underline{\tilde{f}}_l \underline{\tilde{f}}_l^H \mathbf{P}^T \underline{\theta}(n) = \underline{\tilde{\theta}}^T \underline{\tilde{f}}_l \underline{\tilde{f}}_l^H \underline{\tilde{\theta}} \leq 1 \quad \forall\, l = 0, \ldots, \lfloor L/2 \rfloor$$

  where $\underline{\tilde{f}}_l \in \mathbb{C}^L$ is the lth column of the DFT matrix $\mathbf{U} \in \mathbb{C}^{L \times L}$, i.e.

$$\underline{\tilde{f}}_l^H = \begin{bmatrix} 1 & e^{-j\omega_l} & \ldots & e^{-j\omega_l (L-1)} \end{bmatrix}$$

Constrained Recursive Estimation (I)

- Consider the time-variant signal model

$$\underline{x}(n) = \mathbf{S}(n)\underline{\theta}_0(n) + \underline{z}(n), \quad n \geq 0$$

where it is assumed that $\underline{\theta}_0(n) \in \Theta \subset \mathbb{R}^M$

- Motivated by the WLS estimator, we reformulate this signal model into

$$\underline{x}_n = \mathbf{S}_n\underline{\theta}(n) + \underline{z}_n, \quad \underline{\theta}(n) \in \Theta$$

where $\underline{x}_n$, $\mathbf{S}_n$ and $\underline{z}_n$ are stacked versions of $\underline{x}(i)$, $\mathbf{S}(i)$ and $\underline{z}(i)$, i.e.

$$\begin{aligned}
\underline{x}_n &= \begin{bmatrix}\underline{x}(n)^T & \cdots & \underline{x}(0)^T\end{bmatrix}^T \\
\mathbf{S}_n &= \begin{bmatrix}\mathbf{S}(n)^T & \cdots & \mathbf{S}(0)^T\end{bmatrix}^T, \\
\underline{z}_n &= \begin{bmatrix}\underline{z}(n)^T & \beta^{-1/2}\underline{z}(n-1)^T & \cdots & \beta^{-n/2}\underline{z}(0)^T\end{bmatrix}^T.
\end{aligned}$$

- Noise $\underline{z}(n)$ Gaussian and temporally uncorrelated, i.e. $\underline{z}_n \sim \mathcal{N}(\underline{0}, \mathbf{C}_n)$ with

$$\mathbf{C}_n = \operatorname{diag}\left(\mathbf{C}(n), \beta^{-1}\mathbf{C}(n-1), \cdots, \beta^{-n}\mathbf{C}(0)\right)$$

Constrained Recursive Estimation (II)

- How can we avoid estimators with growing computational complexity?
  $\Rightarrow$ use concept of sufficient statistics

- A sufficient statistic for our signal model is given by

$$\underline{t}(\underline{x}_n) = \left(\mathbf{S}_n^T \mathbf{C}_n^{-1} \mathbf{S}_n\right)^{-1} \mathbf{S}_n^T \mathbf{C}_n^{-1} \underline{x}_n = \underline{\hat{\theta}}_{\mathsf{WLS}}(n)$$

- Efficiently computed by recursive weighted least squares algorithm (RWLS)

  - in each time step, it updates

$$\begin{aligned} \underline{\hat{\theta}}_{\mathsf{WLS}}(n) &= \underline{t}(\underline{x}_n) & \ldots \quad & \textit{sufficient statistic} \\ \mathbf{R}_n^{-1} &= \left(\mathbf{S}_n^T \mathbf{C}_n^{-1} \mathbf{S}_n\right)^{-1} & \ldots \quad & \textit{inverse correlaton matrix} \end{aligned}$$

$\Rightarrow$ both quantities are needed by the three recursive estimators
$\Rightarrow$ estimators are based on RWLS and use constraint to improve performance

Recursive Constrained ML (RCML)

- The RCML is given by

$$\underline{\hat{\theta}}_{\text{RCML}}(n) = \arg\max_{\underline{\theta}} \; p(\underline{x}_n|\underline{\theta}) \quad \text{s.t.} \quad \underline{\theta} \in \Theta$$

$$= \arg\min_{\underline{\theta}} \; (\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))^T \mathbf{R}_n (\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n)) \quad \text{s.t.} \quad \underline{\theta} \in \Theta$$

At each time step, we have to check $\underline{\hat{\theta}}_{\text{WLS}}(n) \in \Theta$:

if yes: $\underline{\hat{\theta}}_{\text{RCML}}(n) = \underline{\hat{\theta}}_{\text{WLS}}(n)$

if no: find minimum of $(\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))^T \mathbf{R}_n (\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))$ on boundary of Θ

- For our RIR Tracking problem:
  - Constraint is given by

$$|\Theta(\omega_l, n)|^2 = \underline{\theta}(n)^T \mathbf{P} \underline{\tilde{f}}_l \underline{\tilde{f}}_l^H \mathbf{P}^T \underline{\theta}(n) \leq 1 \quad \forall\, l = 0, \ldots, \lfloor L/2 \rfloor$$

$\Rightarrow$ Quadratically constrained quadratic program (QCQP)
$\Rightarrow$ can be efficiently solved

Recursive Affine Minimax (RAMX)

- The RAMX is given by

$$\underline{\hat{\theta}}_{\text{RAMX}}(n) = (\mathbf{I} + \mathbf{M}(n))\,\underline{\hat{\theta}}_{\text{WLS}}(n) + \underline{u}(n)$$

where $\mathbf{M}(n)$ and $\underline{u}(n)$ are the solution to the minimax problem

$$\min_{\mathbf{M},\underline{u}} \max_{\underline{\theta}\in\Theta} \quad \|\mathbf{M}\underline{\theta} + \underline{u}\|^2 + \operatorname{tr}\left\{(\mathbf{I}+\mathbf{M})\mathbf{R}_n^{-1}(\mathbf{I}+\mathbf{M})^T\right\}$$

and the inverse correlation matrix $\mathbf{R}_n^{-1} = (\mathbf{S}_n^T\mathbf{C}_n^{-1}\mathbf{S}_n)^{-1}$ is updated by RWLS

- For our RIR Tracking problem:
  - reduce computational complexity by using $\underline{u}(n) = \underline{0}$ and $\mathbf{M}(n) = \alpha\mathbf{I}$
  - use S-procedure on Θ to reformulate optimization problem into SDP

Recursive Minimum Mean Squared Error

- RMMSE recasts frequentist problem into Bayesian one:
  - use uniform prior on Θ
  - can be motivated by "maximum entropy principle"
- Using Bayesian sufficient statistics for our signal model, RMMSE is given by

$$\underline{\hat{\theta}}_{\text{RMMSE}}(n) = \frac{\int_{\Theta} \underline{\theta} \exp\{-\frac{1}{2}(\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))^T \mathbf{R}_n^{-1}(\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))\} d\underline{\theta}}{\int_{\Theta} \exp\{-\frac{1}{2}(\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))^T \mathbf{R}_n^{-1}(\underline{\theta} - \underline{\hat{\theta}}_{\text{WLS}}(n))\} d\underline{\theta}}$$

- Integrals are solved using Monte Carlo integration:
  - Sampling from a posterior density using rejection sampling
  - Posterior density is truncated Gaussian density (support given by Θ)
- For our RIR Tracking problem:
  - use rejection sampling with $|\mathit{\Theta}(\omega_l, n)|^2 \leq 1$

Simulation Results (I)

- Task: Estimate RIR from moving source to fixed microphone

- Setup:
  - Source
    - Moves along a straight line, 10cm increment between two pos.
    - Overall 11 positions, i.e. source moves one meter in total
    - Source signal: Gaussian noise of length $S = 100$ samples
  - Room:
    - Room size: $3 \times 3 \times 2.3$m, $T_{60} = 120$ms
    - Image source model ($f_s = 12$kHz) results in RIRs with $M = 1241$ taps
  - Estimators:
    - For all three estimators: DFT length is $L = 2^{14} = 16\,384$
    - $I = 3\,000$ samples for RMMSE (MC approx. of integrals)

Simulation Results (II)

- Definitions:
  - Signal-to-noise ratio: SNR $= 10\log_{10}\frac{\|\mathbf{\Theta}(n)\underline{s}(n)\|^2}{(M+S-1)\sigma^2}$
  - normalized error measure: $E = \frac{1}{N}\sum_{n=1}^{N}\frac{\|\hat{\underline{\theta}}(n)-\underline{\theta}(n)\|^2}{\|\underline{\theta}(n)\|^2}$
- Instantaneous estimators:

SNR	WLS	CML	AMX	MMSE
5 dB	1.23$\times 10^{0}$	1.20$\times 10^{0}$	1.21$\times 10^{0}$	1.02$\times 10^{0}$
10 dB	3.88$\times 10^{-1}$	3.88$\times 10^{-1}$	3.87$\times 10^{-1}$	3.72$\times 10^{-1}$

- Recursive estimators: $\beta_{\text{opt}} \approx 0.55$ for SNR $= 5$dB; $\beta_{\text{opt}} \approx 0.28$ for SNR $= 10$dB

SNR	RWLS	RCML	RAMX	RMMSE
5 dB	7.02$\times 10^{-1}$	6.97$\times 10^{-1}$	6.99$\times 10^{-1}$	6.52$\times 10^{-1}$
10 dB	2.98$\times 10^{-1}$	2.98$\times 10^{-1}$	2.98$\times 10^{-1}$	2.94$\times 10^{-1}$

Summary

- Recursive RIR estimation with energy conservation constraint was studied
  - Constraint helps to improve performance
  - RMMSE estimator with uniform prior performs best

- To further improve performance: incorporate additional a priori information

Summary

- Recursive RIR estimation with energy conservation constraint was studied
  - Constraint helps to improve performance
  - RMMSE estimator with uniform prior performs best

- To further improve performance: incorporate additional a priori information

Thank you for your attention

Backup

Frequency Domain Representation

- The condition $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n) - \mathbf{I} \preceq \mathbf{0}$ requires all eigenvalues of $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)$ to be smaller or equal to 1
- Let

$$r(n) = \begin{cases} \sum_{m=0}^{M-1-|n|} \theta_m(n)\theta_{|n|+m}(n) & |n| \leq M-1 \\ 0 & \text{otherwise} \end{cases}$$

denote the (unnormalized) auto-correlation function of the RIR $\underline{\theta}(n)$

- Then, the matrix $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n) \in \mathbb{R}^{S\times S}$ can be written as a symmetric Toeplitz matrix where the first row is given by $r(n)$ for $n = 0, \ldots, S-1$, i.e. $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)$ is bandlimited
- Using the asymptotic equivalence of the eigenvalues of a bandlimited Toeplitz matrix and the corresponding circulant matrix, it follows

$$|\Theta(\omega, n)|^2 = \left| \sum_{m=0}^{M-1} \theta_m(n) e^{-j\omega m} \right|^2 \leq 1 \;\forall\, \omega \in [0, 2\pi)$$

Affine Minimax Estimator for RIR

- **Epigraphic form of the AMX optimization problem**:

$$\min_{\mathbf{M},\underline{u},\tau} \ \tau$$

subject to

$$\begin{bmatrix} \underline{\theta} \\ 1 \end{bmatrix}^T \begin{bmatrix} -\mathbf{M}^T\mathbf{M} & -\mathbf{M}^T\underline{u} \\ -\underline{u}^T\mathbf{M} & \tau - \mathrm{tr}\left\{(\mathbf{I}+\mathbf{M})\mathbf{R}_n^{-1}(\mathbf{I}+\mathbf{M})^T\right\} - \underline{u}^T\underline{u} \end{bmatrix} \begin{bmatrix} \underline{\theta} \\ 1 \end{bmatrix} \geq 0 \ \forall \ \underline{\theta} \in \Theta$$

- **S-procedure**: sufficient condition for the statement

for all $\underline{z}$: $\underline{z}^T\mathbf{F}_0\underline{z} \geq 0$, . . . , $\underline{z}^T\mathbf{F}_{\tilde{L}}\underline{z} \geq 0 \Rightarrow \underline{z}^T\mathbf{G}\underline{z} \geq 0$

to be true is the existence of $\lambda_0, \ldots, \lambda_{\tilde{L}} \geq 0$ such that $\mathbf{G} \succeq \lambda_0\mathbf{F}_0 + \cdots + \lambda_{\tilde{L}}\mathbf{F}_{\tilde{L}}$.

Affine Minimax Estimator for RIR (II)

- **Final optimization problem**:

$$\min_{\substack{\lambda_0 \geq 0, \ldots, \lambda_{\tilde{L}} \geq 0 \\ \tau, \alpha, x}} \tau$$

subject to

$$\begin{aligned} \lambda_l \geq x/L \quad (l = 0, \ldots, \tilde{L}), \\ \begin{bmatrix} x & \alpha \\ \alpha & 1 \end{bmatrix} \succeq \mathbf{0}, \\ \tau - (1 + 2\alpha + x)\,\mathrm{tr}\{\mathbf{R}_n^{-1}\} \geq \lambda_0 + \cdots + \lambda_{\tilde{L}}. \end{aligned}$$

Bayesian Sufficient Statistics

- **Definition**: (introduced by Kolmogorov)
  - $\underline{t}(\underline{x})$ is Bayesian sufficient statistic if for all priors $p(\underline{\theta})$ the a posteriori density $p(\underline{\theta}|\underline{x})$ can be written as $p(\underline{\theta}|\underline{x}) = p(\underline{\theta}|\underline{t}(\underline{x}))$.

    $\Rightarrow$ Given $\underline{t}(\underline{x})$, no additional information can be infered from $\underline{x}$

- Concept directly related to classical sufficiency:
  - using Fisher-Neyman factorization theorem: Classical sufficiency implies Bayesian sufficiency

$$p(\underline{\theta}|\underline{x}) = \frac{p(\underline{\theta})p(\underline{x}|\underline{\theta})}{\int p(\underline{\theta})p(\underline{x}|\underline{\theta})d\theta} = \frac{p(\underline{\theta})g(\underline{t}(\underline{x}),\underline{\theta})}{\int p(\underline{\theta})g(\underline{t}(\underline{x}),\underline{\theta})d\theta} = p(\underline{\theta}|\underline{t}(\underline{x}))$$

  - converse is also true under some regularity conditions