Embedded Optimization

Moritz Diehl
Optimization in Engineering Center OPTEC
and Electrical Engineering Department ESAT
K.U. Leuven,
Belgium

Joint work with
B. Defraene, H.J. Ferreau,
J. Swevers, M. Moonen, J. De Schutter, T. Van Waterschoot,
Overview

- Linear filters vs. embedded optimization
- Which maps can be represented by parametric convex programming (PCP)?
- Application: Optimal clipping for hearing aids
We are interested in maps from one space of sequences:

\[ \ldots, y_{i-1}, y_i, y_{i+1}, \ldots \]

into another:

\[ \ldots, u_{i-1}, u_i, u_{i+1}, \ldots \]

Important special case:

**Linear Time Invariant, Finite Impulse Response (LTI - FIR) Filters**

\[ u_k = \sum_{i=0}^{N} a_i y_{k-i} = Ax \]

where we introduce the shorthand

\[ x = (y_{k-N}, \ldots, y_k) \]

ie., filter output is linear combination of past inputs.
Linear Filters are Everywhere…

In audio processing:
- Dolby
- active noise cancelling
- echo and other sound effects

In control:
- Kalman filter
- PID
- LQR

…but they often need lots online tuning to deal with constraint saturations, gain changes etc.
Alternative: Embedded Optimization

- Idea: obtain **NON-LINEAR** "filter" by solving repeatedly a parametric optimization problem:

\[ u = \arg \min_u g(u, x) \quad \text{s.t.} \quad (u, x) \in \Gamma \]

Examples:
- every LTI FIR Filter can be trivially obtained as:

\[ u = \arg \min_u \left\| u - Ax \right\|_2^2 \]

- A nonlinear map can e.g. be obtained by parametric LP or QP:

\[ u = \arg \min_u \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} g \\ h \end{bmatrix} \quad \text{s.t.} \quad Au + Bx \leq b \]
Embedded Optimization = CPU Intensive, Nonlinear Map

$$u = \arg \min_u \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} g \\ h \end{bmatrix} \text{ s.t. } Au + Bx \leq b$$

$$x \mapsto u = \mu(x)$$
Overview

- Linear filters vs. embedded optimization

- **Which maps can be represented by parametric convex programming (PCP)?**

- Application: Optimal clipping for hearing aids
The ubiquity of convex optimization

THEOREM [Baes, D., Necoara 2008]:

Every continuous map

\[ \mu : \mathbb{R}^{nx} \to \mathbb{R}^{nu} \]

\[ x \mapsto u = \mu(x) \]

can be represented as a **Parametric Convex Program (PCP)**

\[ \mu(x) = \arg \min_u g(u, x) \text{ s.t. } (u, x) \in \Gamma \]

(PCP = objective and constraints convex in variables and parameters, e.g. parametric LP or QP)
Sketch of Proof

Given: graph of \( \mu(x) \)

Construct epigraph \( E \) of \( g(u, x) \):
1) "lift" graph of \( \mu(x) \) using strictly convex \( g^0(x) \) and add "upward" rays
2) take convex hull. Can show that position of minima is conserved

\[
S := \{(x, \mu(x), t) : x \in \Omega, g^0(x) \leq t\}, \\
E := \text{conv}(S)
\]
Overview

- Linear filters vs. embedded optimization
- Which maps can be represented by parametric convex programming (PCP)?
- **Application: Optimal clipping for hearing aids**
Optimal Clipping: Problem formulation

Input audio frame  
Output audio frame

\[
\min_{y \in \mathbb{R}^N} \frac{1}{2} \sum_{i=0}^{N-1} w_i |Y(e^{j\omega_i}) - X(e^{j\omega_i})|^2 \quad \text{s.t.} \quad l \leq y \leq u
\]

'Perceptual difference'  'Clipping constraint'

Weights \( w_i \): perceptual weighting function in FREQUENCY SPACE!

\[
\min_{y \in \mathbb{R}^N} \frac{1}{2} y^T D^T W D y + (D^T W D x)^T y \quad \text{s.t.} \quad h_i(y) = U - y_i \geq 0  \\
\quad h_{i+N}(y) = y_i - L \geq 0 \quad i = 1, \ldots, N
\]

Convex QP in standard form WITH DENSE HESSIAN!
Fast Solution via FFT, Dualization, and External active set strategy

- Formulate dual optimization problem

\[
\lambda^* = \arg \min_{\lambda \in \mathbb{R}^{2N}} \frac{1}{2} \lambda^T B^T \underbrace{H^{-1} B}_\text{Hessian } \tilde{H} \lambda + \underbrace{(C^T e - B^T x)}_\text{Gradient } \tilde{y}^T \lambda \quad \text{s.t. } \lambda \geq 0 \tag{1}
\]

\[
y^* = x - H^{-1} B \lambda^*
\]

- Efficient solution using external active set strategy

1. Check which inequality constraints are violated in the previous solution iterate. In case no inequality constraints are violated, the algorithm terminates.

2. Add these violated constraints to an active set \(S\) of constraints to watch.

3. Solve a small-scale QP corresponding to (1) with those \(\lambda_k(i)\) not in \(S\) set to zero. Evaluation of eq. (2) yields the new solution iterate.
violated constraint indices = nonzero multipliers in small scale dual QP
External Active Set Strategy: 1\textsuperscript{st} Iteration Result
External Active Set Strategy: 1st Iteration Result

add the very few newly violated constraint indices to dual QP, solve again
External Active Set Strategy: 2\textsuperscript{nd} Iteration = Solution
Result: online optimization within 8.7 ms

- overlapping time frames of 8.7 ms each, 512 samples

- **dense Hessian** QP with 512 variables, 1024 constraints

- solution time per time frame with dual and FFT based external active set strategy: below 8.7 ms (if less than 30 from 1024 constraints are active)

Alternatives:
- interior point methods
- gradient projection schemes. In particular Nesterov's optimal scheme shows excellent results, as gradient can be computed extremely fast.
Audio Test Example: Hard Clipped Signal
Audio Test Example: Optimally Clipped Signal
Summary

- Embedded Optimization promises to revolutionize all aspects of control engineering and signal processing
- It needs sophisticated numerical methods
- OPTEC develops open source software for embedded optimization
- Powerful tool in applications:
  - Optimal Clipping: Fourier based external active set strategy (100 Hz)
- Lots of exciting applications in engineering that need ultra-fast real-time optimization algorithms