A fast algorithm for high-order sparse linear prediction

T.L. Jensen[1], D. Giacobello[2], T. van Waterschoot[3], M.G. Christensen[4]

[1] Signal and Information Processing, Dept. of Electronic Systems, Aalborg University, Denmark
[2] Codec Technologies R&D, DTS Inc., Calabasas, CA, USA
[3] Department of Electrical Engineering (ESAT-STADIUS/ETC), KU Leuven, Belgium
[4] Audio Analysis Lab., AD:MT, Aalborg University, Denmark

tlj@es.aau.dk, giacobello@ieee.org, toon.vanwaterschoot@esat.kuleuven.be, mgc@create.aau.dk

Motivation

- Inherent limitations of standard linear prediction (LP) methods.
- Better signal model and meaningful signal representation for speech achievable by introducing sparsity for the predictor and residual.
- Many applications would require real-time algorithms, possible for embedded implementation. What is the method of choice?

Background

- LP provides a compact representation for the signal $x[t]$:

$$x[t] = \sum_{n=1}^{N} \alpha_n x[t-n] + r[t] , \qquad (1)$$

  - $\alpha = [\alpha_n]_{n=1}^N$ are the prediction coefficients,
  - $r[t]$ is the prediction error.
- A common route for estimation of α is via:

$$\underset{\alpha}{\text{minimize}} \; \|x - X\alpha\|_p^p , \qquad (2)$$

  - vectorized version of (1) over a certain frame,
  - $\|\cdot\|_p$ is the p-norm,
  - With $p = 2$: $\alpha^\star = (X^TX)^{-1}X^Tx$.
- Alternative: capture short-term and long-term redundancies jointly.
- Avoid overfitting when increasing the order of the predictor using a sparsity promoting penalty.
- 1-norm penalty on residual offers better modeling for speech.
- A possible convex formulation:

$$\underset{\alpha}{\text{minimize}} \; \|x - X\alpha\|_1 + \gamma\|\alpha\|_1 . \qquad (3)$$

- Solving (3) is however more "complicated" than traditional 2-norm based LP.
- Matrix structure: dense but X and X^TX are Toeplitz.
- Interior-point methods: Diagonal weighting X^TDX destroys the Toeplitz structure (and does not have low displacement rank) $\rightarrow$ resort to $\mathcal{O}(N^3)$ methods for solving linear systems.
- Consider a method where it is possible to exploit the Toeplitz structure.

The considered method

A straight-out-of-the box alternating direction method of multipliers (ADMM) algorithm:

$$\begin{aligned}
\alpha^{(k+1)} &= (X^TX + \gamma^2 I)^{-1} \begin{bmatrix} \gamma I & -X^T \end{bmatrix} \left(y^{(k)} + \begin{bmatrix} 0 \\ -x \end{bmatrix} - u^{(k)} \right) \\
e^{(k+1)} &= x - X\alpha^{(k+1)} \\
y^{(k+1)} &= S_{1/\rho}\left(\begin{bmatrix} \gamma\alpha^{(k+1)} \\ e^{(k+1)} \end{bmatrix} + u^{(k)} \right) \\
u^{(k+1)} &= u^{(k)} + \begin{bmatrix} \gamma\alpha^{(k+1)} \\ e^{(k+1)} \end{bmatrix} - y^{(k+1)} .
\end{aligned}$$

Notice that in this ADMM formulation there is no weighting of X^TX and we may then use fast (and superfast) methods for solving the linear Toeplitz system.

Toeplitz systems

- How to compute $\alpha^{(k+1)}$ efficiently?
- The behavior of $y^{(k)} - u^{(k)}$ from iteration to iteration is currently unclear, so we consider this as a general right-hand side problem

$$(X^TX + \gamma^2 I)\alpha^{(k+1)} = v^{(k)} . \qquad (4)$$

- Using the autocorrelation method, $X^TX + \gamma^2 I = \{t_{|i-j|}\}_{i,j=1}^N = T$ is a symmetric positive definite Toeplitz matrix.
- Levinson algorithm: time complexity $\mathcal{O}(N^2)$.
- Iterative algorithm: same coefficient matrix and changing right-hand side $\rightarrow$ caching.
- Gohberg-Semencul representation:

$$\delta_{N-1} T^{-1} = T_1 T_1^T - T_0^T T_0 \qquad (5)$$

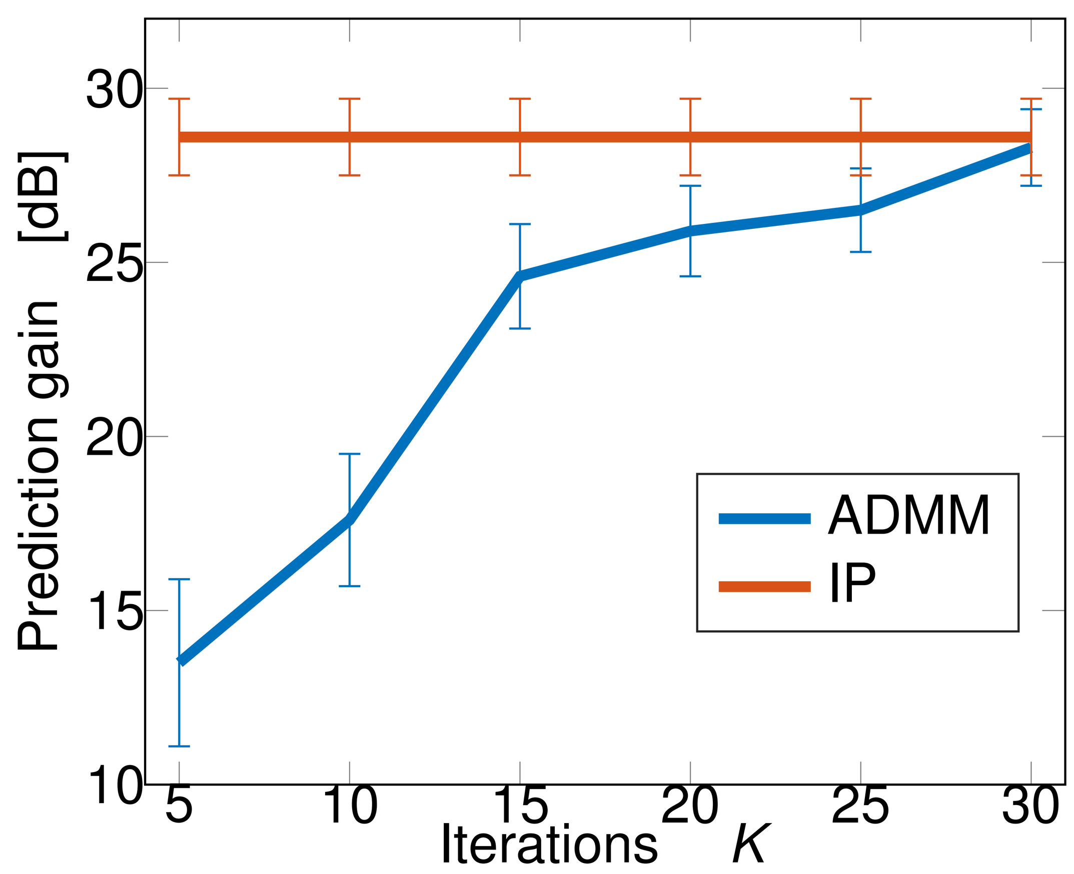
$$T_0 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \rho_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{N-2} & \cdots & \rho_0 & 0 \end{bmatrix} , \qquad (6)$$

$$T_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \rho_{N-2} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \rho_0 & \cdots & \rho_{N-2} & 1 \end{bmatrix} . \qquad (7)$$

- The variables δ_{N-1} and $\rho_0, \dots, \rho_{N-2}$ can be obtained using the Szegö recursion in $\mathcal{O}(N^2)$.
- T_0 and T_1 are Toeplitz: matrix-vector product can be evaluated using FFTs/IFFTs $\rightarrow$ solve system in $\mathcal{O}(N \log N)$.
- Superfast: substitute the Szegö recursion with an $\mathcal{O}(N \log^2 N)$ algorithm. Trade-off at $N = 256$.

Numerical simulations

- Processed the vowel and semivowel phones from the TIMIT database (f_s = 16kHz): 3696 sentences, 462 speakers.
- At least 640 samples $\rightarrow$ total of about 40,000 voiced speech frames.
- Investigate prediction gain as a function of the number of iterations.
- Pick 50 frames randomly, rest used for L-curve analysis to obtain γ.



- Average prediction gains for a fixed number of iterations for the ADMM solution. A 95% confidence interval is shown. The IP solutions is independent of K but shown for the ease of comparison.
- At $K = 30$ the mean value of the IP solution falls within the 95% confidence interval of the ADMM solution $\rightarrow$ the two algorithms exhibit statistically the same performance.

Timings

- C++ implementation + FFTW3 + MKL. Linear regression for $k \in \{5, 10, 15, 20, 25, 30\}$
- Levinson:

$$t_k \approx 7\cdot 10^{-6} + 159\cdot 10^{-6} k \quad [\mathrm{s}], \qquad C^2 = 0.999 .$$

- GS: Gohberg-Semencul + Szegö recursion:

$$t_k \approx 62\cdot 10^{-6} + 55\cdot 10^{-6} k \quad [\mathrm{s}], \qquad C^2 = 0.999 .$$

- Hand-tailored IP method:

$$t_k \approx 1.2\cdot 10^{-3} + 2.0\cdot 10^{-3} k \quad [\mathrm{s}], \qquad C^2 = 0.999 .$$

References

[1] T. L. Jensen, D. Giacobello, T. van Waterschoot, and M. G. Christensen, "Fast algorithms for high-order sparse linear prediction with applications to speech processing," *Speech Comm.*, vol. 76, pp. 143–156, 2016.

[2] D. Giacobello, M. G. Christensen, M. N. Murthi, S. H. Jensen, and M. Moonen, "Sparse linear prediction and its applications to speech processing," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 5, pp. 1644–1657, 2012.

[3] J. R. Jain, "An efficient algorithm for a large Toeplitz set of linear equations," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, no. 6, pp. 612–615, 1979.

[4] G. S. Ammar and W. B. Gragg, "Superfast solution of real positive definite Toeplitz systems," *SIAM J. Matrix Analysis. Appl.*, vol. 9, no. 1, pp. 61–76, 1988.