Increasing accuracy of frequency estimation by decimation

Maciej Niedźwiecki, Tomasz Kłaput and Piotr Kaczmarek
Faculty of Electronics, Telecommunications and Computer Science
Department of Automatic Control
Technical University of Gdańsk
ul. Narutowicza 11/12, Gdańsk , Poland
telex 512302 plg pl fax (48 58) 415821

ABSTRACT

The problem of estimation of a frequency of a complex sinusoidal signal buried in white measurement noise is considered. It is shown that, for a fixed model order, the variance of the autoregressive frequency estimates can be significantly reduced if the signal is decimated prior to modeling. To obtain comparable results by means of processing the signal in the original time scale, one may be forced to use autoregressive models of very high orders. Therefore the major advantage of decimation is dramatic reduction in computation.

1 Introduction

Consider the problem of estimation of a frequency of a complex sinusoidal signal buried in white measurement noise. The problem of frequency estimation is archetypical. There is a large body of statistical literature devoted to comparison of various spectrum estimation techniques in terms of their accuracy and/or resolution – see e.g. [1], [3], [4], [6]. The aim of this paper is different. We focus on a particular parametric frequency estimation technique and investigate how its accuracy is affected by the time scale adopted for the underlying signal model.

The paper can be considered a continuation of the work of Quirk and Liu [5]. Quirk and Liu examined the autoregressive spectrum estimators and showed that their resolution (i.e. the ability to separate neighboring frequencies) can be substantially increased, in the preselected frequency bands, by means of signal decimation. The analysis presented in [5] is carried out for a 'theoretical' autoregressive spectrum, i.e. the one obtained for a *known* sequence of true autocorrelation coefficients of the analyzed signal. Even though all findings are well supported by the results of finite sample experiments, they allow one to infer about the spectral resolution problem only – the accuracy issues are not considered. Since decimation increases the signal to noise ratio but, at the same time, decreases the number of the available samples, its net effect on the estimation accuracy is not obvious. The paper clarifies this issue. Even though the analysis is restricted to a relatively simple case (one complex sinusoid buried in white additive measurement noise) the obtained results seem to provide insights into a more general and rarely discussed problem in system identification.

2 Frequency estimation

Suppose that the available data, obtained by uniform sampling of a continuous-time signal, consists of N samples

$$x(t) = Ae^{j\omega_0 t} + n(t), \quad t = 1, \dots, N \tag{1}$$

where t denotes the normalized discrete time – a dimensionless multiplier of the sampling interval T_s, $\omega_0 = 2\pi T_s f_0$, $|\omega_0| \leq \pi$, denotes the normalized digital angular frequency, and A is the complex amplitude. We will assume, that $n(t) = n_1(t) + jn_2(t) \sim \mathcal{N}(0, \sigma_n^2)$ is an i.i.d. complex Gaussian measurement noise with independent and identically distributed real and imaginary parts $(\mathrm{E}[n_1^2(t)] = \mathrm{E}[n_2^2(t)])$.

To estimate the unknown frequency ω_0 we will build the first-order autoregressive (AR) model of $x(t)$

$$x(t) = ax(t-1) + e(t), \tag{2}$$

where $e(t)$ denotes the modeling error. The unknown autoregressive coefficient a can be estimated from the available data using the method of least squares

$$\begin{aligned} \widehat{a}(N) &= \arg\min_a \sum_{t=2}^{N} |x(t) - ax(t-1)|^2 \\ &= \frac{\sum_{t=2}^{N} x(t)x^*(t-1)}{\sum_{t=2}^{N} |x(t-1)|^2} \end{aligned} \tag{3}$$

and the estimate of ω_0 can be obtained by examining the angular location of the pole associated with the AR model

$$\widehat{\omega}(N) = \arg\, \widehat{a}(N) = \arg\, r(N) \tag{4}$$

where

$$r(N) = \sum_{t=2}^{N} x(t)x^*(t-1)$$

$$= \sum_{t=2}^{N} (Ae^{j\omega_0 t} + n(t))(A^* e^{-j\omega_0(t-1)} + n^*(t-1)). \tag{5}$$

Such choice of $\widehat{\omega}(N)$ is equivalent to maximization of the power spectral density corresponding to the AR model. The method described above can be also interpreted as the first-order Prony's approach – see Section 4 for more comments.

Let $p(N) = \mathrm{Re}\{r(N)\}$ and $q(N) = \mathrm{Im}\{r(N)\}$. Observe that $\omega_0 = \arg r_0$ where $r_0 = \mathrm{E}[r(N)] = p_0 + jq_0 = (N-1)|A|^2 e^{j\omega_0}$. The following approximation will be used to estimate the mean square frequency estimation error

$$\widehat{\omega}(N) - \omega_0 = \arctan \frac{p_0 + \Delta p}{q_0 + \Delta q} - \arctan \frac{p_0}{q_0}$$

$$\cong \frac{q_0\, \Delta p - p_0\, \Delta q}{(1 + (p_0/q_0)^2)(q_0 + \Delta q) q_0} \cong \frac{q_0\, \Delta p - p_0\, \Delta q}{p_0^2 + q_0^2} \qquad (6)$$

where $\Delta p = p(N) - p_0$ and $\Delta q = q(N) - q_0$. The second transition in (6) stems from the Taylor series approximation

$$\arctan x - \arctan x_0 \cong \frac{x - x_0}{1 + x_0^2}$$

and the last transition is the result of neglecting the term Δq in the denominator.

Using (6) one arrives at the following expression (see Appendix)

$$\mathrm{E}[(\widehat{\omega}(N) - \omega_0)^2] \cong \frac{1}{2\xi^2(N-1)} + \frac{1}{\xi(N-1)^2} \qquad (7)$$

where $\xi = |A|^2/\sigma_n^2$ is the signal to noise ratio. This expression will serve as a starting point for our analysis of the benefits of signal decimation.

3 The benefits of signal decimation

In many applications frequency search can be limited to a relatively narrow frequency band known *a priori*. For example, in the mobile radio transmission system considered in [2], analysis of the physical constraints, such as the carrier frequency, the bit rate and the maximum vehicle speed, allow one to restrict the possible range of Doppler shifts, characterizing the channel's time variation, to a certain interval $[-\omega_{\max}, \omega_{\max}]$, where $\omega_{\max} \ll \pi$ (e.g. $\omega_{\max} = \pi/30$). In some other cases the admissible frequency range takes the form $[\omega_{\min}, \omega_{\max}]$, where $\omega_{\max} - \omega_{\min} \ll \pi$. To simplify further considerations we will restrict analysis to the lowpass case only, but all results can be easily generalized to bandpass limitations.

Suppose that the estimated frequency is limited from above $|\omega| < \omega_{\max}$ and denote by d the integer-valued decimation rate. The decimated time series $\widetilde{x}(t)$ can be obtained by taking every d-th sample of the lowpass-filtered input series (for convenience it was assumed that the number of available samples N is an integer multiply of d)

$$\widetilde{x}(t) = x_f(td), \quad t = 1, \ldots, \widetilde{N}, \quad \widetilde{N} = N/d \qquad (8)$$

where

$$x_f(t) = \mathcal{L}[x(t); \pi/d], \qquad (9)$$

and $\mathcal{L}[\cdot; \omega_c]$ denotes lowpass filtering with cutoff frequency ω_c. Note that when $d < d_{\max} = \pi/\omega_{\max}$, the explored frequency band remains within the frequency range of the decimated signal. The filter preserves the spectrum inside the band of interest and suppresses the aliasing caused by subsequent down-sampling by d.

3.1 Ideal filtration

First of all, we will consider the case where $\mathcal{L}$ is an ideal lowpass filter. Such filter preserves without changes the sinusoidal component of $x(t)$ and reduces by the factor of d the variance of its noise component. Consequently, in a new time scale (1) should be replaced with

$$\widetilde{x}(t) = A e^{j\widetilde{\omega}_0 t} + \widetilde{n}(t), \quad t = 1, \ldots, \widetilde{N} \qquad (10)$$

where $\widetilde{\omega}_0 = \omega_0 d$ and $\widetilde{n}(t) \sim \mathcal{N}(0, \sigma_n^2/d)$.

We note that even though the filtered noise $n_f(t) = \mathcal{L}[n(t); \pi/d]$ is correlated, its down sampled version $\widetilde{n}(t) = n_f(td)$ is, similarly as the original noise sequence $n(t)$, white. This means that the results derived in Section 2 are still applicable, provided that the number of data points N is replaced with $\widetilde{N} = N/d$ and the signal to noise ratio ξ is replaced with $\widetilde{\xi} = \xi d$. Additionally, since (7) is a large sample approximation, we need to assume that $\widetilde{N} \gg 1$ i.e. $d \ll N$.

Denote by $\widehat{\widetilde{\omega}}(N)$ the estimate of $\widetilde{\omega}_0$, obtained by means of processing the decimated time series and by

$$\widehat{\widehat{\omega}}(N) = \widehat{\widetilde{\omega}}(N)/d$$

the corresponding estimate of ω_0. Straightforward calculations based on (7) lead to

$$\mathrm{E}[(\widehat{\widehat{\omega}}(N) - \omega_0)^2] \cong \frac{1}{2\xi^2 d^3 (N-d)} + \frac{1}{\xi d (N-d)^2}\ . \qquad (11)$$

Comparison of (7) and (11) (note that for $d = 1$ both expressions become identical) shows clearly advantages of signal decimation. Assuming that $N \gg d$, the decimation based approach allows one to reduce the variance of the frequency estimate at least d times, i.e. to increase estimation accuracy by the factor of at least $\sqrt{d}$.

3.2 Nonideal filtration

Since the ideal lowpass filter cannot be realized in practice, the results derived in the previous subsection can serve only as an indication of the limiting bound. To check how filter nonidealities affect accuracy of the frequency estimates, consider a very simple form of lowpass filtering – signal averaging

$$x_f(t) = \frac{1}{d} \sum_{i=0}^{d-1} x(t-i). \qquad (12)$$

Note that the decimation scheme which incorporates (12), corresponds to a very simple form of the time scale modification, known as data aggregation – each consecutive group of d signal samples is replaced with one, averaged sample.

When averaging is used, the model of the decimated signal becomes

$$\widetilde{x}(t) = \widetilde{A}e^{j\widetilde{\omega}_0 t} + \widetilde{n}(t), \quad t = 1, \ldots, \widetilde{N} \tag{13}$$

where $\widetilde{A} = A(1+e^{-j\omega_0 d})/d(1+e^{-j\omega_0})$, and the remaining quantities are defined analogously as in the previous subsection (in particular, note that aggregation preserves whiteness of the measurement noise). The corresponding expression for the estimation variance is

$$\mathrm{E}[(\widehat{\omega}(N) - \omega_0)^2] \cong \frac{1}{2\xi_d^2 d^3(N-d)} + \frac{1}{\xi_d d(N-d)^2} \tag{14}$$

where

$$\xi_d = \xi \left(\frac{\sin(\omega_0 d/2)}{d\,\sin(\omega_0/2)} \right)^2$$

denotes the modified signal to noise ratio. Since for $\omega_0 > 0$ it holds $\xi_d < \xi$, the estimation variance is larger than that obtained for the idealized scheme – cf. (11). However, if the attenuation of the sinusoidal component of $x(t)$, caused by nonideal filtering, is not significant, signal decimation allows one to considerably improve estimation results.

4 Simulation results

Several simulation experiments were carried out to verify the results of theoretical analysis, as well as to check some more general research hypotheses, which are difficult to confirm analytically. The examined time series consisted of 2000 samples of a complex sinusoidal signal ($A = 1.0$, $\omega_o = \pi/50$) buried in a complex white Gaussian noise ($\sigma_n^2 = 0.1$, SNR=10dB).

Figure 1 (left plot) shows the experimental error curve – dependence of the mean square frequency estimation error on the decimation rate d – obtained for the first-order autoregressive modeling with data aggregation. All results were averaged over 200 simulation runs, corresponding to 200 independent realizations of the noise sequence. For comparison, two theoretical error curves are shown in the same figure: one for the method based on data aggregation (10) and the other one for the method which incorporates ideal decimation (14). Note a very good agreement of the experimental results with the corresponding theoretical evaluations.

Figure 1 (right plot) shows comparison of the error plots, obtained for the first-order autoregressive model, in the case where data averaging is combined with down-sampling (the aggregation approach) and in the case of averaging without subsequent down-sampling. In the second case the the signal is analyzed in the original time scale. Even though in both cases averaging increases the signal to noise ratio by the same amount, the approach based on data aggregation yields considerably better results. This shows clearly that the accuracy improvements achieved by signal decimation cannot be

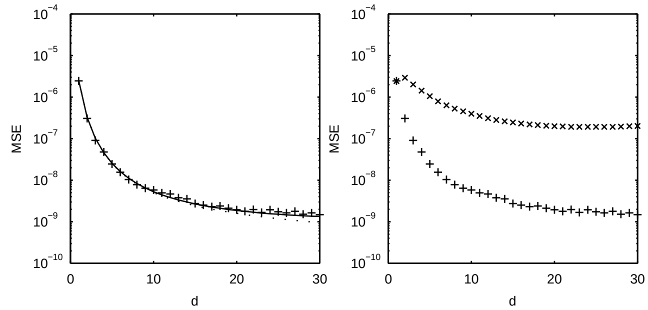

Figure 1: Comparison of the mean square frequency estimation errors obtained under data aggregation (+) with the theoretical error curves derived for aggregation (solid line) and for the ideal decimation (dotted line). The right plot shows comparison of the mean square frequency estimation errors obtained for data aggregation (+) and for data averaging without down-sampling (x). All results correspond to the autoregresive model of order 1.

attributed only to the increase in SNR caused by low-pass filtering (which is an obvious, and hence somewhat trivial, effect). The change of the time scale is another important factor which allows one to further decrease the estimation variance.

To improve accuracy of frequency estimates in the presence of measurement noise, one should increase the order of the autoregressive model beyond the number of sinusoidal components. The approach based on factorization of the overestimated AR model is known as the least squares Prony's method [3]. The estimation principles remain unchanged – the frequency estimates are evaluated as phase angles of complex roots of the characteristic polynomial associated with the AR model. However, since the number of roots exceeds the number of sinusoids, all roots must be divided to signal-related roots, which are further used for frequency estimation, and noise-related roots, which are neglected. Inclusion of the extraneous, noise-related roots at the identification stage allows for better positioning of the signal-related roots, and hence increases accuracy of the frequency estimates.

Even for a single complex sinusoid buried in additive noise, statistical analysis of the properties of the high-order Prony's method is too difficult to perform. For this reason we will refer to the simulation evidence. Results of computer simulations, summarized in Fig. 2 and Fig. 3, shed more light on the problem of frequency estimation.

First, it seems that decimation is beneficial irrespective of the order of the applied AR model. The error curve obtained for the 5th order Prony's method with signal decimation (see Fig. 2) resembles the analogous curve obtained for the first-order model. As expected, raising the order of autoregression allows one to increase accuracy of the frequency estimates. Signal decimation yields even better results.

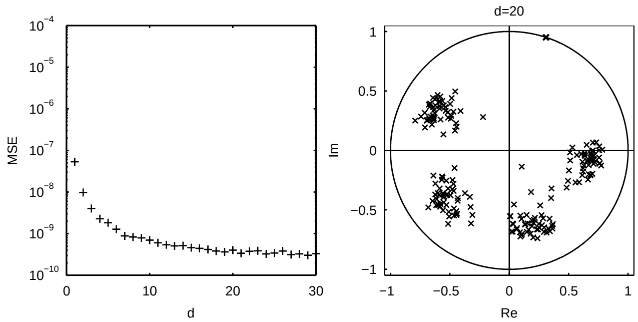

Figure 2: Estimation results obtained under data aggregation (+) for the AR model of order 5. The right plot shows the roots of the characteristic polynomial obtained in 50 trials using the least squares Prony's method. The roots are shown relative to the unit circle.

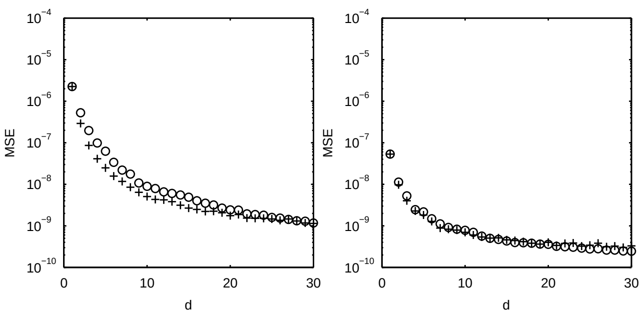

Figure 3: Comparison of the estimation results obtained with data aggregation (+) and without data aggregation (o). The model orders are equal to 1 and d (left plot) and 5 and $5d$ (right plot), respectively.

The second observation is more fundamental and tackles the model order - time scale equivalence problem. It is natural to expect that if the analog process is modeled in two different time scales, one d times smaller than the other, the model orders should stay in the same proportions for the results to be equivalent, namely for fast sampling the order of the model should be d times larger than for slow sampling. Figure 3 compares results obtained with signal decimation for the 1st order and 5th order Prony's method, with the analogous results obtained without signal decimation for the models of orders d and $5d$, respectively. Since the corresponding plots are almost identical, it is clear that decimation does not yield any 'absolute' accuracy improvements – the low-order model offers similar performance as the high-order model obtained for an undecimated signal. The major advantage of decimation is reduction in computation. Since parametric frequency estimation requires factorization of polynomials, even for small values of d the computational savings can be significant. Additionally, one avoids technical problems caused by the fact that with growing model order it becomes increasingly difficult to correctly separate the signal-related roots of the characteristic polynomial from the noise-related roots.

Appendix

According to (6)

$$\mathrm{E}[(\widehat{\omega}(N)-\omega_0)^2] \cong \frac{\mathrm{E}[(\Delta p)^2]q_0^2 + \mathrm{E}[(\Delta q)^2]p_0^2 - 2\mathrm{E}[\Delta p \Delta q]p_0 q_0}{(p_0^2+q_0^2)^2} \ . \qquad (15)$$

Using the orthogonality relationships $\mathrm{E}[n(t)n(s)] = \mathrm{E}[n^*(t)n^*(s)] = 0, \ \forall t,s$ and $\mathrm{E}[n(t)n^*(s)] = \delta(t-s)\sigma_n^2$, where $\delta(t)$ denotes the Kronecker delta function, one obtains (after straightforward but tedious calculations)

$$\mathrm{E}[|\Delta r|^2] = \mathrm{E}[|\Delta p\,|^2] + \mathrm{E}[|\Delta q|^2] = 2(N-1)\sigma_n^2|A|^2 + (N-1)\sigma_n^4 \qquad (16)$$

and

$$\mathrm{E}[(\Delta r)^2] = \mathrm{E}[(\Delta p)^2] - \mathrm{E}[(\Delta q)^2] + 2j\mathrm{E}[\Delta p \Delta q] = 2(N-2)\sigma_n^2|A|^2 e^{2j\omega_0} \ . \qquad (17)$$

Solving (16) and (17) with respect to $\mathrm{E}[(\Delta p)^2]$, $\mathrm{E}[(\Delta q)^2]$ and $\mathrm{E}[\Delta p \Delta q]$ one obtains

$$\begin{aligned}\mathrm{E}[(\Delta p)^2] &= (N-1)\sigma_n^2|A|^2 + (N-1)\sigma_n^4/2 \\ &+ (N-2)\sigma_n^2|A|^2\cos 2\omega_0 \ , \\ \mathrm{E}[(\Delta q)^2] &= (N-1)\sigma_n^2|A|^2 + (N-1)\sigma_n^4/2 \\ &- (N-2)\sigma_n^2|A|^2\cos 2\omega_0 \\ \mathrm{E}[\Delta p \Delta q] &= (N-2)\sigma_n^2|A|^2\sin 2\omega_0 \ .\end{aligned}$$

Finally, after substituting the above expressions into (15) and noting that $p_0^2 + q_0^2 = |r_0|^2 = (N-1)^2|A|^4$, one arrives at (7).

References

[1] Kay S.M. (1988). *Modern Spectral Estimation*. Prentice Hall.

[2] Kłaput T. and Niedźwiecki M. (2001). "A novel approach to estimation of Doppler frequencies of a time-varying communication channel", submitted for publication.

[3] Kumaresan R. (1993). "Spectral analysis", in *Handbook for Digital Signal Processing*, Mitra S.J. and Kaiser J.F Eds., pp. 1143-1242.

[4] Marple L. (1987). *Digital Spectral Analysis with Applications*. Prentice Hall.

[5] Quirk M.P. and Liu B. (1983). "Improving resolution for autoregressive spectral estimation by decimation", *IEEE Trans. Acoust., Speech, Signal Process.*, pp. 630-637.

[6] Stoica P. and Moses R.L. (1997). *Introduction to Spectral Analysis*. Prentice Hall.