Local Regression with Meaningful Parameters

Rafael A. Irizarry [*]

Abstract

Local regression or loess has become a popular method for smoothing scatterplots and for nonparametric regression in general. The final result is a "smoothed" version of the data. In order to obtain the value of the smooth estimate associated with a given covariate a polynomial, usually a line, is fitted locally using weighted least squares. In this paper we will present a version of local regression that fits more general parametric functions. In certain cases, the fitted parameters may be interpreted in some way and we call them meaningful parameters. Examples showing how this procedure is useful for signal processing, physiological, and financial data are included.

KEY WORDS: Local Regression, Harmonic Model, Meaningful Parameters, Sound Analysis, Circadian Pattern.

1 Introduction

Local regression estimation is a method for smoothing scatterplots, $(\mathbf{x}_i, y_i)$ for $i = 1, \dots, n$, in which the fitted value at, say, $\mathbf{x}_0$ is the value of a polynomial fit to the data using weighted least squares where the weight given to $(\mathbf{x}_i, y_i)$ is related to the distance between $\mathbf{x}_i$ and $\mathbf{x}_0$ (Cleveland 1979; Cleveland and Devlin 1988). Some theoretical work exists, for example Stone (1977) shows that estimates obtained using the local regression methods have desirable asymptotic properties. Recently, Fan (1992,1993) has studied minimax properties of local linear regression. In this paper we will

[*]Rafael A. Irizarry is Assistant Professor, Department of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, MD 21205 (E-mail: rafa@jhu.edu). The author would like to thank David Wessel, Adrian Freed, George Katinas, and Heather Bell for providing the data. The author would also like to thank the editor and the referee for suggestions that have improved this paper. Partial support for this work was provided by the Johns Hopkins University School of Public Health Faculty Innovation Fund.

concentrate on the conceptual aspects.

Through the exploration of data representing sound signals produced by musical instruments, we present a version of local regression, which fits harmonic functions instead of polynomials. Local estimates of the frequency and amplitudes of harmonic components are obtained as well as a smooth version of the data. The fitted parameters, frequency and amplitudes, may have meaningful interpretations. When this is the case we will say that they are *meaningful parameters*. We explore the possibility of applying this procedure to other data that appear to be approximately periodic. Finally, we discuss the possibility of extending this version of local regression to other, more general, parametric functions where the parameters may be interpreted in some specific way. In the next section we examine an example where the fitted parameters contain useful information.

Sounds, data, and software related to the examples described in this paper can be obtained from the author's web page: `http://biosun01.biostat.jhsph.edu/~ririzarr`

2 Local Regression

Local regression or loess is a smoothing technique designed to accommodate data for which we assume the observed y_i's are the outcome of a random process defined by:

$$Y_i = f(\mathbf{x}_i) + \epsilon_i, i = 1, \dots, n$$

with $f(\mathbf{x})$ a "smooth" function of a covariate $\mathbf{x}$ and ϵ_i an independent identically distributed random error. For the examples presented in this paper the covariate is time so we will use the notation $\mathbf{x}_i = t_i$ to denote the time associated with the ith measurement y_i.

Loess is a numerical algorithm that prescribes how to compute an estimate $\hat{f}_0$ of $f(t_0)$ for a specific t_0. By repeating the procedure for all $t_0 = t_1, \dots, t_n$, we obtain a *smooth*, $\hat{f}_i, i = 1, \dots, n$. How do we compute an estimate of $f(t_0)$? Given t_0 we define a weight for each $t_i, i = 1, \dots, n$ using $w_i(x_0) = W(|t_i - t_0|, h)$, with W a non-negative function decreasing with Δ and $W(\Delta, h) = 0$ for $\Delta \geq h$. The statistical package S-Plus uses $W(\Delta, h) = \{1 - (\Delta/h)^3\}^3$ for $\Delta < h$ (Cleveland et al. 1993). The span or window size h controls the "smoothness" of the final estimate (Cleveland 1979; Cleveland and Devlin 1988).

Once the weights have been defined we find the $\hat{\beta}$ that minimizes

$$\sum_{i=1}^{n} w_i(t_0)\{y_i - s(t_i;\beta)\}^2 \tag{1}$$

with $s(t;\beta)$ some polynomial, usually a line (S-Plus has two options, a line or a parabola). We estimate $f(t_0)$ with $\hat{f}_0 = s(x_0, \hat{\beta})$.

In Figure 1 we see an example taken from Diggle, Liang, and Zeger (1994) . The figure shows a scatter diagram of CD4 cell counts versus time since seroconversion for men with HIV. The first plot shows the fitted line when $h = 1$ year, for $t_0 = -2, 0$, and 2 years. The second plot shows the smooth $\hat{f}_i, i = 1, \dots, n$, the final result. Notice that this last plot shows a useful summary of the data: on average CD4 counts go down right after seroconversion and are close to being fixed before and after. However, in this particular case, the fitted slopes and intercepts contained in the $\hat{\beta}$s are not used to convey this information.

3 Local Harmonic Regression

Many time series data are modeled with the so-called signal plus noise models. Many methods have been proposed to analyze these types of data, see for example Brockwell and Davis (1996). In many cases the signal is the component of interest. Non-parametric methods, such as loess, have been used to obtain useful estimates of the signal in such models. In this section we examine a specific example where the local behavior of the signal can be well approximated with a harmonic function. This suggests that we use harmonic functions instead of polynomials for $s(t,\beta)$ in equation (1). We call this procedure local harmonic regression or *lohess*.

3.1 Musical Sound Signals

Sound can be represented as a real-valued function of time. This function can be sampled at a small enough rate so that the resulting discrete version is a good approximation of the continuous one. This permits one to study musical sounds as discrete time series. Physical modeling (Fletcher and Rossing 1991) suggests that many musical instruments'

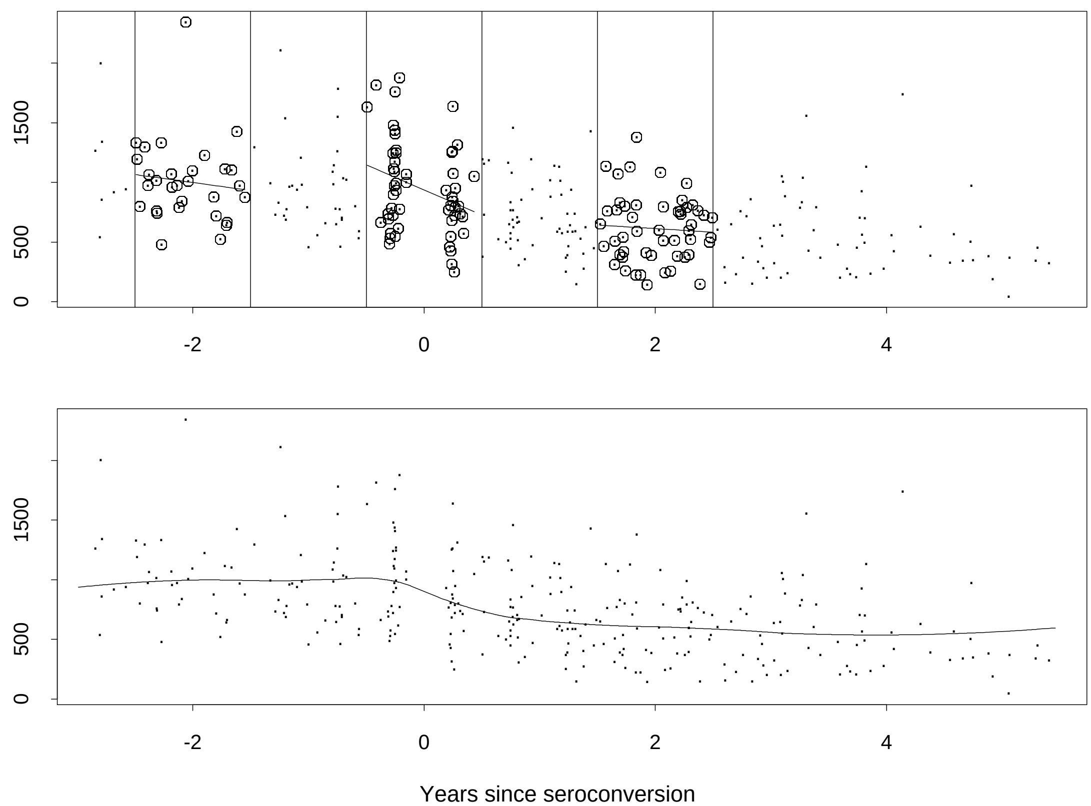

Figure 1: CD4 cell count since seroconversion for HIV infected men.

sounds may be characterized by a approximately periodic signal plus stochastic noise model,

$$Y_i = f(t_i) + \epsilon_i, i = 1, \dots, n.$$

The time series is regular, so without loss of generality we will assume it has a one second duration and set $t_i = i/n$. Researchers are interested in separating these two elements of the sound and finding parametric representations with musical meaning (Rodet 1997).

For many instruments one may argue that the local behavior of the signal produced when playing, say, a C-note is periodic, see Pierce (1992) for details. We may verify this from the data by computing and examining *spectrograms*. For data $y_i, i = 1, \dots, n$, we define the spectrogram at time $t_0 = i_0/n$ with

$$I(t_0, \lambda) = \frac{1}{2\pi(2M+1)} \left[\left\{ \sum_{i=i_0-M}^{i_0+M} \cos(2\pi\lambda\, i/n)\, y_i \right\}^2 + \left\{ \sum_{i=i_0-M}^{i_0+M} \sin(2\pi\lambda\, i/n)\, y_i \right\}^2 \right].$$

Here $(2M+1)/n$ is some suitable window size. For any i_0, if the data $y_{i_0-M}, \ldots, y_{i_0+M}$ is periodic with fundamental frequency λ_0 cycles per second the function $I(t_0, \lambda)$ will have peaks at the multiples of λ_0 (Bloomfield 1976). As an example, in Figure 2, we show spectrograms of the signals produced by a violin playing a C-note, an oboe playing a C-note, and a guitar playing a D-note, all of them sampled at 44100 measurements per second. Dark shades of grey represent large values. The dark horizontal bands at frequencies that are multiples of 260 Hz. (146 Hz. for the guitar) suggest that locally the signals are periodic with that fundamental frequency. The fact that for different times the darkness of these horizontal bands changes suggests that it is only locally that the signals are periodic. Similar results are obtained for other "harmonic" instruments.

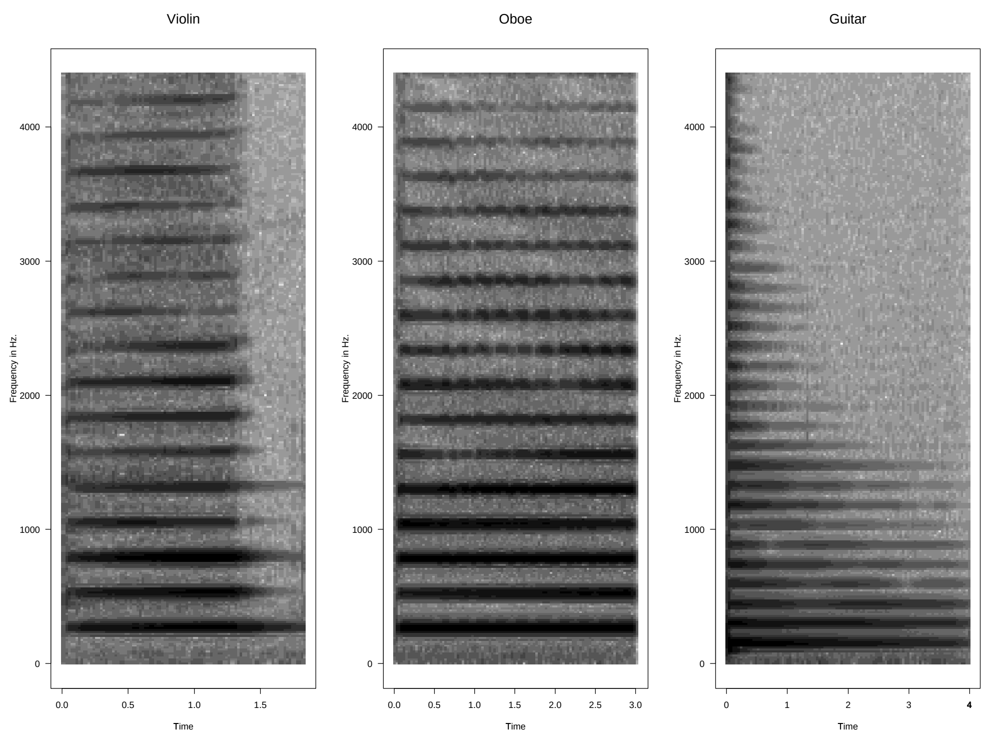

Figure 2: Spectrograms of the sound signals of an oboe, a violin, and a guitar.

Any periodic deterministic regular time series $x_1, \ldots, x_n$ with period p can be written as

$$x_i = a_0 + \sum_{k=1}^{p/2-1} \left\{ a_k \cos\left(\frac{2\pi k}{p}\, i\right) + b_k \sin\left(\frac{2\pi k}{p}\, i\right) \right\} + a_{p/2} \cos(\pi i)$$

for $i = 1, \ldots, n$ (Bloomfield 1976). We define $\lambda = n/p$ as the fundamental frequency and its units are cycles per every n measurements. This suggests that for time series data that appear to be locally periodic we may use lohess, with

$$s(t_i,\beta) = \mu + \sum_{k=1}^{K} \left\{ a_k \cos\left(\frac{2\pi k}{n}\lambda\, i\right) + b_k \sin\left(\frac{2\pi k}{n}\lambda\, i\right)\right\} \tag{2}$$

$$= \mu + \sum_{k=1}^{K} \rho_k \cos\left(\frac{2\pi k}{n}\lambda\, i + \phi_k\right) \tag{3}$$

in equation (1). In this case the parameter we fit is $\beta = (\mu, \lambda, \rho_1, \ldots, \rho_K, \phi_1, \ldots, \phi_K)'$ with $\rho_k = \sqrt{a_k^2 + b_k^2}$ and $\phi_k = \arctan(-b_k/a_k)$ for $k = 1, \ldots, K$, and $K \leq n/(2\lambda)$ fixed. The parameters in β are useful for interpretation. They are meaningful parameters as will be described in the next section.

3.2 Meaningful Parameters

As described in Section 2, when using loess we compute a different $\hat{\beta}$ for each $i = 1, \ldots, n$. We use the subscript $\hat{\beta}_i, i = 1, \ldots, n$ to denote the different fitted values. In the case of loess, these fitted values can be used to give, for example, local slope estimates. However, in general, the smooth $\hat{f}$ itself conveys more information about the data than the fitted parameters $\hat{\beta}_i$'s. With lohess the fitted parameters $\hat{\beta}_i$ contain what could be interpreted as the local mean level $\hat{\mu}_i$, and local estimates of the fundamental frequency $\hat{\lambda}_i$ and amplitudes $\hat{\rho}_{k,i}, k = 1, \ldots, K$ and phases $\hat{\phi}_{k,i}, k = 1, \ldots, K$ of each sinusoidal component or harmonic. Another useful quantity is

$$\left(\sum_{k=1}^{K} \hat{\rho}_{k,i}^2\right)^{1/2},$$

which we refer to as the total amplitude. In the case of lohess, the fitted parameters may sometimes be considered more informative than the smooth $\hat{f}$. For some musical applications, the fitted parameters are the only quantities of interest and are used as parametric representations of sound. Timbre analysis, sound recreation, time-scale/pitch modification, and timbre morphing/modification are some examples of applications that use such parametric representations. See Irizarry (1998) for more details.

In order to give the reader a better idea of how these applications work and how they are useful in practice, we describe a timbre morphing example. Timbre morphing is the process of combining two or more sounds to create a new sound with "intermediate" timbre. Morphing can be used to create interesting sounds that are not found in nature, but that have the characteristics of naturally occurring sounds. An interesting example is the recreation of a castrato voice (Depalle et al. 1995). This was done to produce the sound-track for *Farinelli*, a film about the famous 18th century castrato. To simulate Farinelli's voice, the voice of a counter tenor and a soprano were analyzed with a procedure similar to lohess. The two meaningful parameters were combined in a way that produced a new parametric representation for a timbre similar to that of a castrato.

Now we consider an example of lohess applied to data presented in this paper (and available on the author's web-page). We demonstrate how we can modify the timbre of the oboe sound of Figure 2 using the results from lohess. Timbre modification is the process of changing a sound by manipulation of musically meaningful parametric representations. To obtain a smooth $\hat{f}$ and fitted values of meaningful parameters for the oboe sound, we use lohess with $K = 15$ and $h = 20$ milliseconds. In Figure 3 we see the resulting smooth version of the oboe sound signal, the residuals, and some of the estimated parameters. The fitted fundamental frequency and amplitudes of the first 5 sinusoidal components agree with what we hear: the instrumentalist is playing a vibrato and a tremolo.

The $\hat{\beta}_i$s provide a parametric representation for the harmonic part of the oboe sound, which we can construct using $s(t_i, \hat{\beta}_i)$. If we convert $s(t_i, \hat{\beta}_i)$ into sound we are able to hear a harmonic sound very similar to the original (indistinguishable from the original for most listeners). The residuals have a noisy sound (we must amplify the residuals in order to hear them). It is said that if one listens to an oboe very closely one can hear the voice of a soprano singing at an octave above the fundamental frequency of the oboe signal. We can modify the oboe sound through our parametric representation and bring out the hidden soprano. That is, by creating the sound $s(t_i, m(\hat{\beta}_i))$ where $m(\cdot)$ is a function that modulates the amplitudes of the even harmonics (the even harmonics are related to a sound an octave above the original). See Irizarry (1998) for more details. On the author's web-page there is a music demo where one can hear all the sounds mentioned in this example.

If we perform the same analysis on a 3 second segment of the violin sound, we would then have two parametric representations, say $\hat{\beta}_i^{(oboe)}, \hat{\beta}_i^{(violin)}, i = 1, \ldots, n$. A morph of these two sounds can be created by $s(t_i, m(\hat{\beta}_i^{(oboe)}, \hat{\beta}_i^{(violin)}))$ with $m(\cdot)$ a function that combines the parametric representation in order to obtain a satisfactory morph. In practice

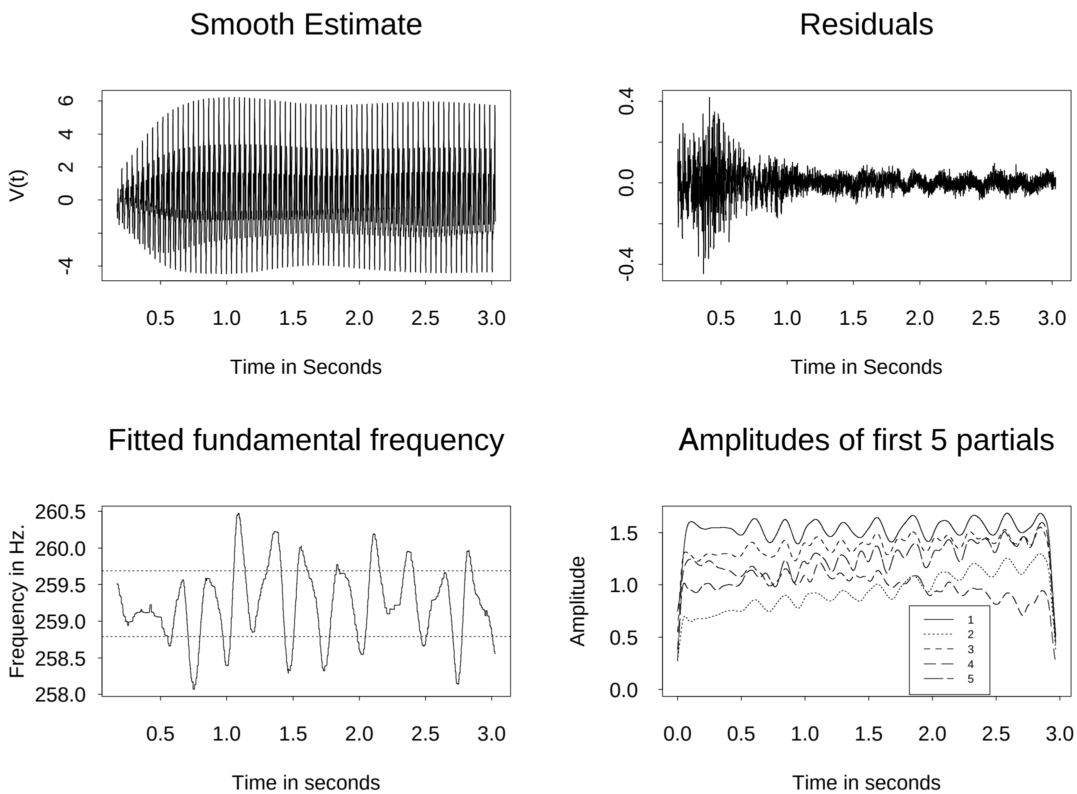

Figure 3: Smooth estimate of an oboe sound signal data with residuals and estimated meaningful parameters.

one must "tweak" to obtain a useful $m(\cdot)$, but in essence it defines new amplitudes by averaging the two originals, that is $\rho_k^{(morph)} = \delta\hat{\rho}_{k,i}^{(oboe)} + (1-\delta)\hat{\rho}_{k,i}^{(violin)}$ with $0 < \delta < 1$ determining if the morph sounds more like the oboe or the violin. See Depalle, García and Rodet (1995) for more details.

3.3 Standard error approximations

Loader (1999) presents finite sample results that may be used to obtain an approximation of the variance of our estimates. Irizarry (2000) presents asymptotic results related to local fitting of harmonic functions. In this section we present the relevant results in a heuristic fashion.

Loader (1999, Chapter 2) presents variance expressions based on finite sample calculations for local regression methods. As pointed out by the referee, the analogous form of the result presented in Loader (1999) for the non-linear problem is

$$\text{cov}(\hat{\beta}_t) = \sigma^2 J_{1,t}^{-1} J_{2,t} J_{1,t}^{-1} \tag{4}$$

where

$$J_{m,t} = \sum_{i=1}^{n} w_i(t)^m \frac{\partial s(t_i,\beta)}{\partial \beta} \frac{\partial s(t_i,\beta)}{\partial \beta}^T .$$

Here we are assuming that the errors are independent and identically distributed with variance σ^2.

For the harmonic regression case, Irizarry (2000) develops some asymptotic results. Consider that a locally periodic signal can be expressed as

$$Y_i = s(t_i; \beta(t_i)) + \epsilon_i, i = 1, \ldots, n$$

with $s(t_i; \beta(t_i))$ as in (2) but with the parameter β now depending on t_i. Irizarry (2000) shows that if 1) the sample rate n is high enough, 2) $\beta(t_i)$ is smooth enough, and 3) the fundamental frequency is large enough, then for any t_0, we can find small estimation windows $(t_0 - h/2, t_0 + h/2)$ with $\beta(t_i)$ well approximated by $\beta(t_0)$ within $(t_0 - h/2, t_0 + h/2)$ and with enough data points nh receiving positive weight so that the estimates obtained with lohess have desirable asymptotic properties, such as asymptotic normality of the estimates. To get an idea of why, consider that conditions 2) and 3) give us that $s(t_i; \beta(t_i))$ is approximately harmonic within the estimation window, condition 1) gives us many data points. We can use the results presented in Irizarry (2000) to obtain approximate marginal standard errors for our parameters. For example, for independent identically distributed errors we have

$$\text{var}(\hat{\lambda}_0) \approx \frac{2\,\hat{\sigma}_0^2}{nh^3} \left\{ \frac{W_0^2 U_2 - 2 W_0 W_1 U_1 + W_1^2 U_0}{(W_0 W_2 - W_1^2)^2} \right\} \left\{ \sum_{k=1}^{K} k^2 \hat{\rho}_{k,i}^2 \right\}^{-1} \text{ and} \tag{5}$$

$$\text{var}(\hat{\rho}_{k,i}) \approx \frac{2\,\hat{\sigma}_0^2}{nh} \left(\frac{U_0}{W_0^2} \right). \tag{6}$$

Here $\hat{\sigma}_0^2$ is a local estimate of the variance of the errors and W_0, W_1, W_2, U_0, U_1, and U_2 are constants defined by

$$W_j = \frac{1}{2h} \int_{-h}^{h} \left(\frac{u}{2h} + \frac{1}{2} \right)^j W(u,h)\, du \quad \text{and} \quad U_j = \frac{1}{2h} \int_{-h}^{h} \left(\frac{u}{2h} + \frac{1}{2} \right)^j W(u,h)^2\, du \text{ for } j = 0, 1, 2.$$

Notice that $u/2h + 1/2$ takes u from $[-h, h]$ to $[0, 1]$.

For the music signals, the sample rate ($n = 44100$ measurements per second) is large, the harmonic parameters can be considered to be slowly varying (as seen in the spectrogram) and the fundamental frequency (for the oboe sound it was was 260 cycles per second) is large enough so that in relatively small windows ($h = 20$ milliseconds) we still have enough oscillation (about 5) to have the harmonic model as a good approximation and enough data points ($n_0 = 882$) to be able to use the large sample approximations. The normality approximation motivated by these asymptotics permits us to construct approximate confidence intervals. However, these confidence intervals should be interpreted with care as it is difficult to assess the assumption that "$\beta(t)$ is smooth enough".

Loader (1999, page 39) points out that asymptotic expansions "should never be considered an alternative to (finite sample results) for actually computing variance". However, computing (4) may be computationally intensive, as we need to perform matrix inversions and multiplications for each $i = 1, \ldots, n$. The asymptotic results presented in Irizarry (2000) provide approximations that are in closed form. Furthermore, one can show that when the number of cycles within the estimation window is large, (5) and (6) are equivalent to the expressions obtained with (4). In our experience with sound signals, when one has more than 5 cycles in the estimation window the results are practically equivalent, for 3-5 cycles they are usefully close. This means that even if we don't use the asymptotic normality result, we may still find (5) and (6) to be useful approximations to the variance estimates obtained from (4). In the software section of the author's web-page there is S-Plus code that compares (4) and (5)-(6).

For many data sets it is not convenient to assume constant variance. For example, for sound signals the noise is usually stronger during the beginning of the note, or what musicians call the attack. This is agreement with the residual plot seen in Figure 3. However, (4), (5) and (6) permit a different estimate of σ^2 for each estimation window, but assume that within the estimation windows the errors are independent and identically distributed. We justify this with a heuristic argument: for our example, we believe the way the variance, $\sigma_i^2 = \text{var}(\epsilon_i), i = 1, \ldots, n$, varies with i is slow enough for us to consider data within an estimation window to be independent and identically distributed. We can use the $\hat{\sigma}_i^2$s as a smooth estimate of the variances σ_i^2. Furthermore, for many examples, especially for time series data, the independence assumption may not be appropriate. The results presented in Irizarry (2000) are developed for locally stationary noise. However, extending the results presented in this paper to more general cases is of interest but is not discussed here.

3.4 Other data sets

Many time series data have trends for which one may argue that locally the behavior is periodic. Different methods have been proposed for estimating this trend, see for example Cleveland et. al. (1990). In situations where the assumptions discussed in the previous section seem approximately true, we propose lohess as a useful alternative. In this section we briefly discuss the application of lohess to a biological data set.

Physiological measurements that vary in a circadian pattern in humans have been extensively studied, see for example Greenhouse, Kass, and Tsay (1987) and Wang and Brown (1996). Systolic blood pressure is an example, it is lowest during sleep and early morning and rises after awakening. Figure 4 shows measurements of systolic blood pressure taken for 208 days with an average of about 48 measurements taken per day on a patient that suffers from a condition dubbed circadian hyper-amplitude-tension (CHAT). This condition is diagnosed, for example, when the patient's daily blood pressure range is larger than what is considered normal. CHAT may occur without an increase of over-all blood pressure, but is believed to put the patient at risk as well. Physicians are interested in how three different treatments affect the patient's condition. From day 1-80 the patient was taking nifedippine, from days 81-125 blocalcin is added to the treatment, and during days 126-208 blocalin is administered during night instead of morning. See Katinas et. al. (1999) for details.

A plot of 15 days of data, also seen in Figure 4, suggests that these measurements in fact follow a circadian pattern. However, one may expect this pattern to change with time. For example, a patient may have smaller fluctuations if a treatment is being effective against CHAT. From just looking at the data it is hard to see if the patient's blood pressure behaves differently in the three periods defined above.

The instrument that was used to take these measurements is a bit inaccurate and measurement error may be present. We use lohess with $K = 3$ and $h = 7$ days to smooth the data and see if any useful information, not seen in the data plot, is obtained. In this case we have prior knowledge of what the fundamental frequency λ should be, namely 1 cycle per day. In situations like this we may choose to set λ to a fixed value instead of estimating it. In this case, lohess is equivalent to computing spectrogram values for $k\lambda, k = 1, \dots, K$. Furthermore, using (2) we can express (3) as a linear model and obtain approximate standard errors in a straight forward fashion.

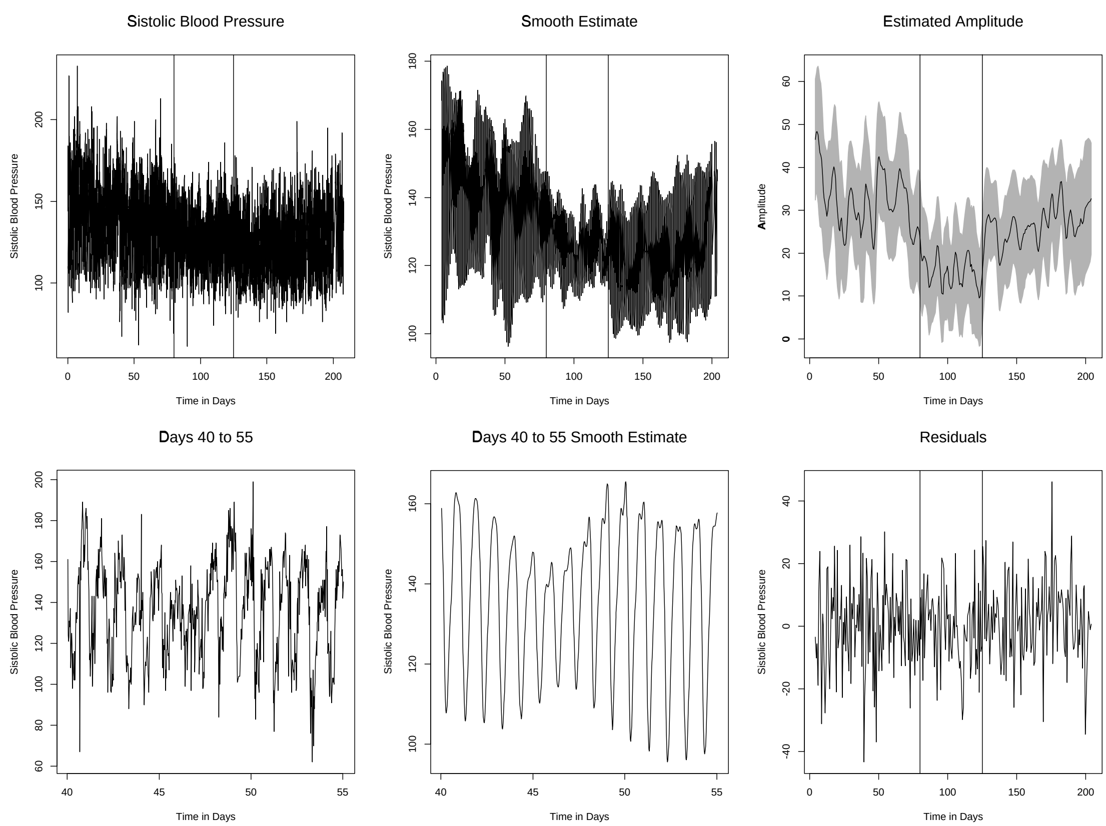

Figure 4: 208 and 15 days of systolic blood pressure measurements taken from a heart patient. Smooth estimate of systolic blood pressure data with residuals and estimated meaningful parameters.

In Figure 4, we see the smooth version and residuals for the the data. The smooth version of the data clearly shows what could be considered to be a circadian pattern. When looking at the smooth $\hat{f}$ we see a drop in mean level at around day 80 and also notice that, between days 80 and 125, there seems to be a drop in total amplitude. After day 125 the amplitude estimates increase slightly. This may be interpreted as evidence that the the second treatment is effective against CHAT. This is seen more clearly in the plots showing the fitted total amplitude. To explore the possibility that these changes are "real", we plot marginal ± 2 standard errors around the fitted amplitudes. The amplitudes plot seems to be the most informative plot in terms of showing how CHAT is being affected by the treatments.

In this case the ± 2 standard errors should not be considered point-wise confidence intervals, since it is hard to argue that the assumptions needed for the asymptotic results are met. For example, we used a window size of 7 days which contains about 7 cycles, but only 280 data points. Within this window the behavior is not as well approximated by a harmonic function as in the music example. If we believe the shape of the oscillations during days that are close together are more similar than for days that are far apart, then using smaller window sizes is an alternative we could consider to improve the local approximation. However, in doing so we would reduce the number of data points and/or number of cycles in the window, making the variance approximations inappropriate. In any case, lohess has been useful in terms of finding that something different is occurring during days 80 to 125.

3.5 Computational Issues

The S-Plus functions written for lohess are available on the software section of the author's web-page. In this section we discuss some of the problems that the users may encounter.

In practice, fitting equation (2) is more convenient than fitting equation (3) because λ is the only non-linear parameter and we can use minimization routines that take advantage of this, see Bates and Lindstrom (1986). The minimization routine used by lohess is the S-Plus function `nlfit` with the `algorithm` option set to `"plinear"`. The function needs starting values for the non-linear parameter. In practice, we have found that it is important that this value is close to the "true" frequency. Notice that every multiple of the true fundamental frequency is a local minimum of (1).

If in (2) $2K + 1$ is not much less than the number of points considered in the local estimation, we are fitting a "saturated model". The final estimate may not be much smoother than the original signal. To avoid this, one possibility is to choose a K, smaller than p/2, that achieves the desired smoothness given a particular span h. In practice one has to balance between K and h to obtain a reasonable estimate. For the examples presented in this paper, appropriate values were found by using our scientific intuition combined with trial and error. See Irizarry (2001) for some data driven procedures for choosing K and h.

4 Extensions

The idea of using non-polynomial local models has been explored by, for example, Hjort and Jones (1996) and Loader (1999). Extensive development of theoretical properties are discussed in their work. In this paper, we have developed practical aspects much more deeply and presented an example in sound analysis where the procedure seems to be successful. We also demonstrated how the fitted parameters have meaningful interpretations in these examples. We believe that this idea can be useful in the analyses of other data sets where functions other than polynomial or harmonic functions may be used for $s(t, \beta)$ in equation (1).

For example, if we look at any 30 year period of the daily closing prices of the DJIA, the data appear to be exponentially growing. However, different periods have different growth rates. By using local regression with a span of 30 years and

$$s(t_i, \beta) = a + b \exp(\gamma t_i)$$

in equation (1), we may obtain a smooth version of this data. In this case the parameters are also meaningful. For example, we can view $\hat{\gamma}_i$ as a measure of average growth rate in the span being considered. In Figure 5 we show the smooth version of the data, the residuals, and fitted meaningful parameters. By looking at the fitted growth rate plot, we can clearly see the recessions of the 30's and 70's, the roaring 20's, and the booming 80's. Furthermore, notice that if we compute the average over time of the smooth yearly growth estimate, we see that it is about 7%-8%, the number given by financial "experts" as what one should expect stocks to grow.

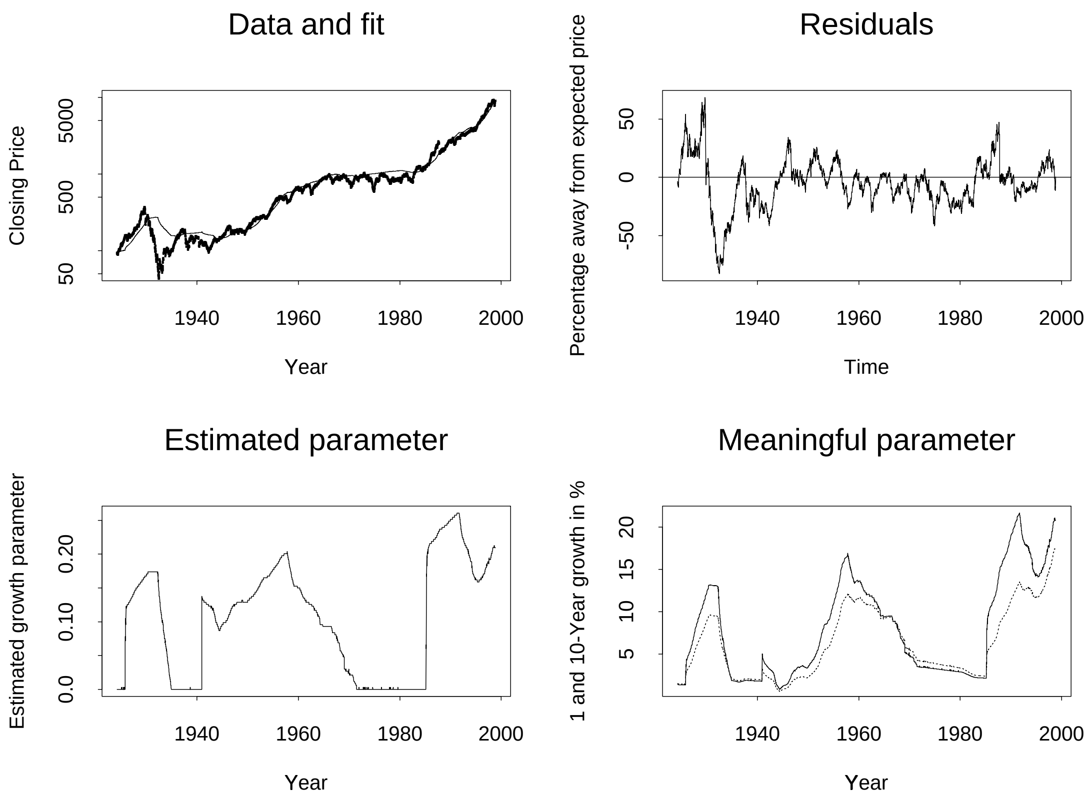

Figure 5: Smooth estimate of Dow Jones industrial average data with residuals and estimated meaningful parameters.

Notice that the results obtained from this analyses may not be very different than those obtained by fitting local lines to a log transform of the data. Taking log transforms is usually considered more useful than the exponential model with additive errors developed here. However, we believe that the analysis presented in this section serves as an example of how loess may be modified and we hope it motivate readers think of how it can be used in other applications.

5 Concluding Remarks

In this paper we have seen how for certain data sets local regression can be modified in order to produce meaningful parameters. The resulting procedure may be used for smoothing data and for obtaining useful information about the local behavior of the data. In particular, we introduced lohess and showed how it can be useful for smoothing time series data with approximately periodic trends. It is important to note that these procedures are intended as exploratory data analysis tools. We do not suggest that the fitted parametric functions are models for the data. However, scientists have found the plots resulting from the procedure presented in this paper insightful, a fact which has led to more detailed analyses.

References

Bates, D. and Lindstrom, M. (1986). Nonlinear least-squares with conditionally linear parameters. *Proceedings, Statistical Computing Section*, pages 152–157.

Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction*. John Wiley and Son, New York.

Brockwell, P. J. and Davis, R. A. (1996). *Introduction to Time Series and Forecasting*. Springer-Verlag.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6:3–33.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.

Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:596–610.

Cleveland, W. S., Grosse, E., and Shyu, W. M. (1993). Local regression models. In Chambers, J. M. and Hastie, T. J., editors, *Statistical Models in S*, chapter 8, pages 309–376. Chapman & Hall, New York.

Depalle, P., García, G., and Rodet, X. (1995). The recreation of a castrato voice, Farinelli's voice. In *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 242–245, New Paltz, NY. IEEE.

Diggle, P. J., Liang, K.-Y., and Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Oxford Science Publications.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Society*, 87:998–1004.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics*, 21:196–216.

Fletcher, N. H. and Rossing, T. D. (1991). *The Physics of Musical Instruments*. Springer-Verlag, New York.

Greenhouse, J. B., Kass, R. E., and Tsay, R. S. (1987). Fitting nonliner models with ARMA errors to biological rhythm data. *Statistics in Medicine*, 6:167–183.

Irizarry, R. A. (2000). Asymptotic distribution of estimates for a time-varying parameter in a harmonic model with multiple fundamentals. *Statistica Sinica*. In press.

Irizarry, R. A. (2001). Information criteria for model selection in local likelihood estimation. *Journal of the American Statistical Association*. To appear.

Katinas, G., Cornelissen, G., Irizarry, R., Schaffer, E., Homans, D., Rhodus, N., Schwartzkopff, O., Siegelova, J., Palat, M., and Halberg, F. (1999). Case report of coexisting elderly mesor-hypertension and circadian amplitude-hypertension (CHAT). *Geronto-Geriatrics*, 2:68–86.

Pierce, J. (1992). *The Science of Musical Sound*. Freeman, New York.

Rodet, X. (1997). Musical sound signals analysis/synthesis: Sinusoidal+residual and elementary waveform models. In *Proceedings of the IEEE Time-Frequency and Time-Scale Workshop (TFTS'97)*, Coventry, UK. IEEE.

Stone, C. J. (1977). Consistent nonparametric regression. *The Annals of Statistics*, 5:595–620.

Wang, Y. and Brown, M. B. (1996). A flexible model for human circadian rhythms. *Biometrics*, 52:588–596.