Electrocardiogram estimation using Lagrange interpolation

Om Prakash Yadav1 and Anil Kumar Sahu, Ph.D Associate Professor
1Department of Electrical and Electronics Engineering, PES Institute of Technology and Management, Shivamogga, Karnataka, 577204, India
2Department of Electronics and Communication Engineering, Bharat Institute of Engineering and Technology, Mangalpally, Hyderabad, Telangana, 501510, India
3E-mail: anilsahu82@gmail.com

An electrocardiogram records activity of cardiac which is collected through the electrodes positioned on specific locations on the human body. These signals are required for cardiac-related issues. Electrocardiogram signals are quasi-stable signals and are often contaminated by artefacts. The magnitude and frequency of these artefacts are different, and thus their effect on electrocardiogram is also variable. These artefacts must reduce to a significantly optimum level to make them interpretable. This paper proposes an algorithm to estimate electrocardiogram signals from the MIT-BIH database signal using Lagrange interpolation through Chebyshev nodes. The performance evaluation of the proposed method has been assessed through standard estimation tools. Performance parameters revealed that the algorithm is suitable for signal estimation irrespective of the characteristic of artefact introduced to the signal.

Introduction: An electrocardiogram (ECG) is an essential tool for related cardiac treatments. The signal is acquired through electrodes placed on the patient’s limbs and chest. One cycle of ECG [1, 2] is normally characterised by P, QRS, T, and U waves of different shapes, amplitudes and frequencies. Common artefacts affecting ECG signals can be obtained from [3]. Since the timing pattern of these waves is diagnostically important [4], estimation algorithms must preserve timing characteristics even after processing.

So, the very first step in ECG analysis is the reduction of the artefacts. Since artefacts introduced in ECG signals are of variable frequency and amplitude, significant artefacts removal cannot be achieved using a single filter. Different types of recursive and non-recursive filters with unacceptably long transient time were widely used to reduce P1I noises [5–7]. Also, the determination of cut-off frequency for these filters were not so easy. Adaptive filters are also found to reduce the artefact of ECG signals with high transient time, especially on the QRS complex to significant levels in [8]. Recursive filters suffering from numerical instability were also found in literature [9]. Baseline wandering is reduced to a significant level using linear and polynomial filtering. Low-frequency artefacts were reduced using median filters in [10]. Wavelet transforms also find application in ECG signal estimation by analysing the signal in time and frequency domain. Various filters using wavelet transform can be found in [11–14]. Selection of threshold and decomposition level is still a challenge. Neural networks and genetic algorithms were also applied to reduce noise from ECG signals [15]. An effective ECG enhancement technique using total variation was proposed in [16].

Here, we propose an estimation of ECG signals using Lagrange interpolation by utilising Chebyshev nodes.

Lagrange interpolation using chebyshev nodes: The nth degree Chebyshev polynomials are a collection of orthogonal polynomials that can be obtained as [17]:

\[ T_n(x) = \cos(n \cos^{-1}(x)) \quad \text{for } n \geq 0. \] (1)

There are \( n + 1 \) nodes in \([-1, 1]\) for nth order Chebyshev polynomial, which can be obtained as

\[ x_k = \cos\left(\frac{2k + 1}{2(n + 1)} \pi\right) \quad 0 \leq k \leq n. \] (2)

The sequence of polynomials \( p_n(x) \) generated by Chebyshev interpolation method converge uniformly to \( f(x) \) in \([-1, 1]\) [17].

A continuously varying function \( f(x) \) in the same interval can be obtained at \( n + 1 \) nodes by a polynomial of degree \( n \). Let \( f(x_k), 0 \leq k \leq N \) be the samples of ECG of length \( N \) in \([-1, 1]\). Then, these samples can be uniquely interpolated by a polynomial of order \( p \) which is \( n \leq N \), that is, \( p(x_k) \approx f(x_k) \) for each \( k \).

Considering the interpolating function as

\[ p_n(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1} + c_nx^n \] (3)
then, we require

\[ c_0 + c_1x_1 + c_2x_1^2 + \cdots + c_{n-1}x_1^{n-1} + c_nx_1^n = f(x_1) \] (4)

In matrix form, (4) can be rewritten as

\[
\begin{bmatrix}
1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & x_0^n \\
1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
1 & x_n & x_n^2 & \cdots & x_n^{n-1} & x_n^n
\end{bmatrix}
\begin{bmatrix}
c_0 \\
c_1 \\
\vdots \\
c_n
\end{bmatrix}
= \begin{bmatrix}
f(x_0) \\
f(x_1) \\
\vdots \\
f(x_n)
\end{bmatrix} \] (5)

The system (4) will have a unique solution when the Vandermonde determinant of (5) on the extreme left should be non-singular [18]. The determinant of (5) is equal to the product of the terms \( (x_i - x_j) \) for \( i > k \), therefore the points \( x_0, \ldots, x_n \) should be distinct for the determinant to be non-zero.

From the interpolated values \( f(x_k), 0 \leq k \leq n \), we can write \( p_n(x) \) of degree \( n \) using Lagrange’s interpolation [19] as

\[ p_n(x) = \sum_{k=0}^{n} l_k^n(x) f(x_k), \] (6)
where \( l_k^n \) are \( (n + 1) \) Lagrange polynomials of degree \( \leq n \).

\[ l_k^n(x) = \prod_{l=0, l \neq k}^{n} \frac{x - x_l}{x_k - x_l}. \] (7)

Let us now consider \( f(x) \) is \( n + 1 \) times continuously differentiable in \([-1, 1]\), then the interpolation error is [20]

\[ f(x) - p_n(x) = \frac{1}{(n + 1)!} \prod_{l=0}^{n} \max_{-1 \leq \xi \leq 1} |f^{(n+1)}(\xi)|, \xi \in [-1, 1], \] (8)

where

\[ E_n(x) = f(x) - p_n(x). \] (9)

We can minimise the product \( \prod (x - x_k) \) by utilising the Chebyshev nodes of \( T_{n+1}(x) \). This would also help to minimise the upper bound for \( |E_n(x)| \). The following theorem gives an estimate of the error.

If the interpolating points are \( x_0, \ldots, x_n \) in the interval \([-1, 1]\), then \( \prod (x - x_k) \) can be easily minimised. The interpolating points also ensure uniform convergence over the nodes of \( T_{n+1}(x) \) [21]. The same can be estimated by the following theorem.

Theorem. If \( f(x) \) is a polynomial passing through \( x_0, x_1, \ldots, x_n \) points, let \( p_n(x) \) be a Lagrange interpolating polynomial to interpolate \( f(x) \). The interpolation points utilised to interpolate \( p_n(x) \) are the \( n + 1 \) nodes of \( T_{n+1}(x) \), given by (2). Then

\[ |E_n(x)| \leq \frac{1}{2(n + 1)!} \max_{-1 \leq \xi \leq 1} |f^{(n+1)}(\xi)| \quad \forall x \in [-1, 1]. \] (10)

The interpolation problem for \( f(x) \) on \([a, b]\) can be transferred to \( f(y) \) on \([-1, 1]\) using

\[ x = \frac{(b - a)y + (a + b)}{2} \] (11)
Accordingly, the Chebyshev nodes in the same interval can be obtained as

\[ y_k = \cos \left( \frac{2k + 1}{2n} \pi \right), \quad 0 \leq k \leq n \]

Thus, the interpolation points in the same interval (11) would be derived as

\[ x_k = \frac{(b - a)y_k + (a + b)}{2}, \quad 0 \leq k \leq n. \tag{12} \]

The \( E_{\text{ck}}(x) \) for the same interval can be obtained from

\[ |E_{\text{ck}}(x)| \leq \frac{1}{2(n + 1)!} \left| \frac{b - a}{2} \right| \max_{\xi \in [-\pi, \pi]} |f^{(n+1)}(\xi)|. \tag{13} \]

Delving deeper into the advantages of using Chebyshev interpolating nodes, we observe that Runge phenomenon does not occur with the effect that the error tends to decrease with the increasing degree of the Lagrange Chebyshev interpolating polynomial, whereas the same may not be true for equally spaced nodes.

The \( \text{th} \) degree interpolating polynomial \( p_n(x) \) in the interval \( x \in [-1, 1] \) can be expressed as [22]

\[ p_n(x) = \sum_{k=0}^{n} c_k T_k(x). \tag{14} \]

where \( c_k \) can be obtained as

\[ c_k = \frac{2}{n+1} \sum_{j=0}^{n} f(x_j) T_k(x_j), \quad k = 0, \ldots, n. \tag{15} \]

Let \( x = \theta, \theta \in [-\pi, \pi] \), then

\[ c_k = \frac{2}{n+1} \sum_{j=0}^{n} f(\cos \theta_j) \cos(k\theta_j), \quad k = 1, \ldots, n \tag{16} \]

with \( \theta_j = \frac{2\pi j}{n+1}. \) Replacing \( f(\cos \theta) \) by a periodic function \( g(\theta) \),

\[ c_k = \frac{2}{n+1} \sum_{j=0}^{n} g(\theta_j) \cos(k\theta_j), \quad k = 1, \ldots, n. \tag{17} \]

Thus, \( c_k \) is the discrete approximation to the Fourier series

\[ c_k = \frac{1}{\pi} \int_{-\pi}^{\pi} g(\theta) \cos(k\theta) \, d\theta. \tag{18} \]

Applying the trapezoidal rule approximation

\[ c_k \approx \frac{1}{\pi} \sum_{j=0}^{n} g \left( \frac{2j+1}{2n+1} \pi \right) \cos \left( \frac{2j+1}{2n+1} \pi \right). \tag{19} \]

which is same as (17) for \( c_k \). Thus, the Chebyshev interpolation can be effectively utilised by replacing the Fourier transform \( c_k \) by discrete Fourier transform \( c_k \).

**Methods and results:** Here Lagrange–Chebyshev interpolants are used to estimate ECG signals. The data used for testing the algorithm is taken from Physionet repository. The records considered in this article are 10 seconds duration consisting of multiple QRS complexes. The sampling rate is fixed to 360 Hz. However, sampling rate may be varied and its effect can also be observed for estimation algorithms. The records utilised has resolution of 11 bits per sample. All the algorithms are performed in MATLAB environment. The performance parameters utilised are signal-to-noise ratio (SNR), Correlation Coefficient, Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Percentage Root Mean Square Difference (PRD).

Let \( y(n) \) be the actual signal and \( x(n) \) be the estimated ECG signal of having length as \( N \). Then, the difference that is, error between the signals is evaluated as

\[ e_p(n) = x(n) - y(n) \tag{20} \]

The assessment tools for the proposed algorithm is then obtained as:

**Mean Absolute Error**

MAE measures mean absolute difference between the actual signal and estimated ECG signal.

\[ \text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |e_p(i)| \tag{21} \]

**Root Mean Square Error**

RMSE calculates root of mean of the square differences between the actual signal and estimated ECG signal. The parameter is utilised in case where there are large variations in the error.

\[ \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} |e_p(i)|^2}{N}} \tag{22} \]

**Percentage Root-Mean-Square Difference**

PRD is calculated by

\[ \text{PRD} = \sqrt{\frac{\sum_{i=1}^{N} |e_p(i)|^2}{\sum_{i=1}^{N} |y(i)|^2}} \tag{23} \]

where \( y \) is the mean of \( y(n) \). Through this parameter, the extent of baseline/DC level reduced in estimated signal can be obtained.

**Signal to Noise Ratio**

SNR calculates signal strength over noise.

\[ \text{SNR(dB)} = 20\log_{10} \left( \frac{\sum_{i=1}^{N} |y(i)|^2}{\sum_{i=1}^{N} |e_p(i)|^2} \right) \tag{24} \]

**Correlation Coefficient**

This parameter determines the degree to which two signals are correlated in their time-domain characteristics. Value near to 1 indicates close resemblance between the actual signal and estimated ECG signal.

\[ \text{CC} = \frac{(\sum_{i=1}^{N} x(i)y(i)) - (\sum_{i=1}^{N} x(i)) (\sum_{i=1}^{N} y(i))}{\sqrt{\left[ (\sum_{i=1}^{N} x(i)^2) - (\sum_{i=1}^{N} x(i))^2 \right] \left[ (\sum_{i=1}^{N} y(i)^2) - (\sum_{i=1}^{N} y(i))^2 \right]}} \tag{25} \]

SNR and CC are positive-oriented parameters and algorithms producing higher values are considered to be better. Other parameters are negatively oriented and algorithms producing lower values are better. Details about the performance parameters can be obtained from [23].

The proposed estimation model constructs \( p_n(x) \) using (6) with \( N \) ECG samples using the Chebyshev nodes obtained from 2. It may happen that the ECG data may not be available at all the Chebyshev nodes, in such case we can utilise the adjacent ECG samples. We continue increasing the order until our error criterion derived in (13) is met. This Lagrange–Chebyshev signal estimation can be summarised as:

1. Fix the order \( n \) of the Chebyshev polynomial and the tolerance \( \varepsilon \) level for the Lagrange–Chebyshev polynomial estimation.
2. Obtain the Chebyshev nodes \( x_j \), using (2) in the interval \([a, b]\).
3. Calculate \( f(x_j) \) using simple linear interpolation through the two adjacent samples around \( x_j \).
4. Construct \( p_n(x) \) using (6) and (7).
5. Find error \( E_p(x) \) in the same interval using (13).
6. If \( E_p(x) > \varepsilon \), then \( n = n + 1 \) and go to step 2.

In order to estimate one cycle of ECG signal, high-order polynomials were required, which increases computation time also. The interpolation errors for this approach was very high. So, to reduce the computation time and error, the whole signal irrespective of the location of P, QRS, T and U waves, divided into sections using the Bottom-up segmentation method. This method is a recursive algorithm which segments the original signal into large number of small length segments. These small
The MAE obtained is up to order of $10^{-6}$ which is very less. Also the RMSEs for all signals are very low, that is, $10^{-4}$. PRD for good estimation algorithms is expected to be less than 5 and is medically accepted. However, for the proposed model, it is even less than 1. CC values are also nearly equal to 1, indicating the close resemblance with the actual signal. The performance parameters may be further improved either by increasing the order of interpolating polynomial or by segmenting the signal into more number of segments at the cost of estimation time.

In [25], various ECG signals were approximated by Chebyshev polynomials of degree 2000, and obtained PRDs varying from 2.74 to 10.64. Jokic et al. [26] estimated ECG signals from the same database using polynomial function by segmenting one cycle of ECG signal into seven segments. The PRD reported for the record 119 is 5.3; however, we could estimate the same record with PRD as 0.0014. Sandryhaila et al. in [27] found Hermite functions more suitable than transformation techniques to estimated QRS complex with a CR of 11 with 25% approximation error.

Table 1. Quality assessment metrics for Lagrange–Chebyshev interpolation

| Signal | MAE × 10^{-6} | RMSE × 10^{-4} | PRD | SNR (dB) | CC  |
|--------|---------------|----------------|-----|----------|-----|
| 10    | 1.8344        | 0.6570         | 0.0019 | 44.8285 | 0.9877 |
| 104   | 2.8941        | 0.9553         | 0.0017 | 43.0592 | 0.9882 |
| 108   | 0.1513        | 0.1116         | 1.7153 × 10^{-4} | 66.9131 | 1.0000 |
| 112   | 0.7953        | 0.2874         | 4.6281 × 10^{-4} | 55.9675 | 0.9999 |
| 115   | 1.8073        | 0.5245         | 0.0030 | 50.3352 | 0.9945 |
| 117   | 2.5893        | 1.3054         | 0.0014 | 43.6540 | 0.9887 |
| 119   | 2.5893        | 1.3054         | 0.0014 | 43.6540 | 0.9887 |
| 122   | 3.9897        | 1.1316         | 0.0023 | 47.6133 | 0.9968 |
| 201   | 0.8445        | 0.3399         | 9.7624 × 10^{-4} | 44.8090 | 0.9994 |
| 205   | 0.7007        | 0.3827         | 1.1696 × 10^{-4} | 53.0546 | 1.0000 |
| 207   | 12.0150       | 2.4682         | 0.0139 | 23.1663 | 0.8837 |
| 214   | 5.7031        | 1.6179         | 0.0090 | 32.2947 | 0.9535 |
| 220   | 3.9218        | 1.4584         | 0.0031 | 41.4559 | 0.9941 |

Fig. 1 A portion of original and Lagrange–Chebyshev interpolated record 214

Conclusion: ECG signal records voltage potentials of cardiac activity. Various types of noises often contaminate these signals. ECG signals are estimated using the Lagrange interpolation method. The Chebyshev nodes are used as interpolation points to increase convergence and accuracy. The results obtained are quite convincing and are superior to those reported in the existing literature. The results can be further improved with increment in order of the Chebyshev polynomial or number of segments. Although the method produces convincing results, order of the polynomial is high. The order can also be further reduced by segmenting the signal into suitable number of segments.

The proposed estimation algorithm is presented in Table 1. The performance parameters may be further improved either by increasing the order of interpolating polynomial or by segmenting the signal into more number of segments at the cost of estimation time.

In [25], various ECG signals were approximated by Chebyshev polynomials of degree 2000, and obtained PRDs varying from 2.74 to 10.64. Jokic et al. [26] estimated ECG signals from the same database using polynomial function by segmenting one cycle of ECG signal into seven segments. The PRD reported for the record 119 is 5.3; however, we could estimate the same record with PRD as 0.0014. Sandryhaila et al. in [27] found Hermite functions more suitable than transformation techniques to estimate QRS complex with a CR of 11 with 25% approximation error.

References

1. Gómez Clapers, J., Casanella, R.: A fast and easy-to-use ecg acquisition and heart rate monitoring system using a wireless steering wheel. IEEE Sens. J. 12(3), 610–616 (2011).
2. Hamdan, M., et al.: Detection of abnormal heart conditions based on characteristics of ECG signals. Measurement 125, 634–644 (2018).
3. Friesen, G.M., et al.: A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans. Biomed. Eng. 37(1), 85–98 (1990).
4. Bharadwaj, A., Kamath, U.: Techniques for accurate ECG signal processing. EE Times 14, 1–7 (2011).
5. Paul, J., et al.: Recursive algorithm for efficient representation of ECG signals. Electron. Lett. 34(6), 526–528 (1998).
6. Rani, S., et al.: Comparative study of FIR and IIR filters for the removal of baseline noises from ECG signal. Int. J. Comput. Sci. Inf. Technol. 2(3), 1105–1110 (2011).
7. Meidani, M., Mashoufi, B.: Introducing new algorithms for realising an FIR filter with less hardware in order to eliminate power line interference from the ECG signal. IET Signal Process. 10(7), 709–716 (2016).
8. Thakor, N.V., Zhu, Y.S.: Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection. IEEE Trans. Biomed. Eng. 38(8), 785–794 (1991).
9. Cuomo, S., et al.: A revised scheme for real time ECG signal denoising based on recursive filtering. Biomed. Signal Process. Control 27, 134–144 (2016).
10. Łęski, J.M., Henzel, N.: ECG baseline wander and powerline interference reduction using nonlinear filter bank. Signal Process. 85(4), 781–793 (2005).
11. Kaur, I., et al.: ECG signal analysis and arrhythmia detection using wavelet transform. J. Inst. Eng. (India) Series B 97(4), 499–507 (2016).
12. Karthikeyan, P., et al.: ECG signal denoising using wavelet thresholding techniques in human stress assessment. Int. J. Electr. Eng. Inf. Technol. 4(2), 306 (2012).
13. Alfaouri, M., Daqrouq, K.: ECG denoising by sparse wavelet shrinkage. Amer. J. Appl. Sci. 9(3), 276–281 (2008).
14. Mankandand, M.S., Dandapat, S.: Effective quality-controlled spilt-based ECG coding strategy under noise environments. Electron. Lett. 44(20), 1182–1184 (2008).
15 Youssif, R.S., Purdy, C.N.: Combining genetic algorithms and neural networks to build a signal pattern classifier. *Neurocomputing* **61**, 39–56 (2004)
16 Ning, X., Selesnick, I.W.: ECG enhancement and QRS detection based on sparse derivatives. *Biomed. Signal Process. Control* **8**(6), 713–723 (2013)
17 Pavliović, V.D., et al.: 1D and 2D economical fir filters generated by Chebyshev polynomials of the first kind. *Int. J. Electron.* **100**(11), 1592–1619 (2013)
18 Björck, K., et al.: Solution of Vandermonde systems of equations. *Math. Comput.* **24**(112), 893–903 (1970)
19 Chan, W.C.C., Chyan, C.J., Srivastava, H.: The lagrange polynomials in several variables. *Integral Transforms Spec. Funct.* **12**(2), 139–148 (2001)
20 Caporale, G.M., Cerrato, M.: Chebyshev polynomial approximation to approximate partial differential equations. *Comput. Econ.* **35**(3), (2010)
21 Yang, W.Y., et al.: *Applied numerical methods using MATLAB*. John Wiley & Sons, Hoboken, NJ (2005)
22 Rababah, A.: Transformation of Chebyshev–Bernstein polynomial basis. *Comput. Methods Appl. Math.* **3**(4), 608–622 (2003)
23 Yadav, O.P., Ray, S.: ECG signal characterization using Lagrange-Chebyshev polynomials. *Radioelectron. Commun. Syst.* **62**(2), 72–85 (2019)
24 Yadav, O.P., Ray, S.: Smoothening and segmentation of ECG signals using total variation denoising–minimization-majorization and bottom-up approach. *Procedia Comput. Sci.* **85**, 483–489 (2016)
25 Tchiotsop, D., Ionita, S.: ECG data communication using Chebyshev polynomial compression methods. *Telecommunications* **11**(2), 22–32 (2010)
26 Jokic, S., et al.: Efficient ECG modeling using polynomial functions. *Elektronika ir Elektrotechnika* **11**(4), 121–124 (2011)
27 Sandryhaila, A., et al.: Efficient compression of QRS complexes using hermite expansion. *IEEE Trans. Signal Process.* **60**(2), 947–955 (2012)