AMATH 563: COMPUTATIONAL REPORT 1

NGHI NGUYEN

Applied Math Department, University of Washington, Seattle, WA
nghiiuu@uw.edu

1. Introduction

In this report, we use the Cosine Series, Monomial, and RBF dictionaries to approximate the Levy function with recorded observation points. The total 50 observation points are the values of the Levy function with noises. To avoid overfitting to the observation points while minimizing the approximation error, we solved a least square problem with varying regularization value and number of basis.

2. Methods

2.1. **Theoretical Background.** For an unknown function $f(x)$, recorded observations of f often will have noise such that the observed points are not on the f curve, rather near the curve. To approximate the unknown function f, first consider the interval $[a, b]$ and divide it into K uniform grids. Denote the collection of grid points as $\{x_k\}_{k=0}^{K-1}$ and the vector of observation points as $y = [y_0, ..., y_{K-1}]^T \in \mathbb{R}^K$ where

$$y_k = f(x_k) + \xi_k$$

for $k = 0, ..., K-1$. Here, ξ_k denotes the noise at the observed point y_k and $\xi_k \sim N(0, \sigma^2)$ where 0 is the mean and σ^2 is the standard deviation of the normal distribution the noise follows.

2.2. **The Approximation of the Levy Function.** In this paper, let $K = 50$ and the unknown function is the Levy function,

$$f(x) = \sin(\pi w)^2 + (w-1)^2(1 + \sin(2\pi w)^2), \quad w = 1 + \frac{10x - 1}{4}$$

for the grid points $\{x_k\}_{k=0}^{49} \in [-1, 1]$ and the observation vector $y = [y_0, ..., y_{49}]^T \in \mathbb{R}^{50}$.

To approximate the Levy Function, consider the combination of N basis functions denoted as $\{\psi_n\}_{n=0}^{N-1}$. Then, the approximation function is as follow,

$$g(x) = \sum_{n=0}^{N-1} c_n \psi_n(x)$$

Date: April 15, 2025.

. To minimize the space between the approximation function as the point x_k, $g(x_k)$, and the corresponding observed point y_k, consider the Least Squares problem,

$$\begin{aligned}\hat{\mathbf{c}} :=& \underset{\mathbf{c}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{2}\sum_{k=0}^{K-1} |g(x_k) - y_k|^2 - \frac{\lambda}{2}||\mathbf{c}||_2^2 \\ =& \underset{\mathbf{c}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{2}\sum_{k=0}^{K-1} |\sum_{n=0}^{N-1} c_n\psi_n(x_k) - y_k|^2 - \frac{\lambda}{2}||\mathbf{c}||_2^2 \\ =& \underset{\mathbf{c}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{2}||\Psi\mathbf{c} - y_k||_2^2 - \frac{\lambda}{2}||\mathbf{c}||_2^2\end{aligned}$$

where Ψ is a KxN matrix where the kth row is the N basis functions at the point x_k and the nth column is the value of the ψ_n function evaluated at the points $\{x_k\}_{k=0}^{K-1}$.

$$\Psi = \begin{bmatrix} \psi_0(x_0) & ... & \psi_{N-1}(x_0) \\ . & & . \\ . & & . \\ . & & . \\ \psi_0(x_{K-1}) & ... & \psi_{N-1}(x_{K-1}) \end{bmatrix}$$

. The second term in the Least Square optimization problem is to decrease overfitting of the observation points since these points have include noise and are not on the Levy function Curve. Denote function within the minima as $L(\mathbf{c})$. Since norms and norms are convex functions, then the argument $\hat{c}$ can be found by solving for $\nabla(L\hat{\mathbf{c}}) = 0$.

$$\begin{aligned}\nabla L(\mathbf{c}) =& \nabla(\frac{1}{2}||\Psi\mathbf{c} - y_k||_2^2 - \frac{\lambda}{2}||\mathbf{c}||_2^2) \\ =& \frac{2}{2}\Psi^T(\Psi\mathbf{c} - y_k) - \frac{2\lambda}{2}\mathbf{c} \\ =& \Psi^T(\Psi\mathbf{c} - y_k) - \lambda\mathbf{c}\end{aligned}$$

Solving for $\hat{c}$, the coefficient vector that minimizes $L(\mathbf{c})$ is

$$\hat{\mathbf{c}} = (\Psi^T\Psi - \lambda\mathbf{I})^{-1}\Psi^T y$$

Thus, the Levy function can be approximated with a sum of combinations of basis function through solving a least square problem. The next subsections discusses different basis functions and variations of N and λ.

2.3. **Basis Functions.** The different basis functions used to approximate the Levy function in this report include: the Cosine, Monomial, Radial Basis Functions (RBF), and the Cosine and Sine basis. The first basis functions are implemented with different values $N = [10, 50, 100]$, the number of combinations of the basis function, and the values $\lambda = [10^{-10}, 10^{-2}, 1]$, the regularization parameter for $\hat{\mathbf{c}}$. The Cosine and Sine basis are implemented $N = [5, 10, 25]$ and $\lambda = [10^{-12}, 1.2]$ for variation. Each dictionary will be examined for a fixed value N with the varying regulation parameter.

Cosine Series: $\psi_n = \cos(n\pi x)$
Monomials $\psi_n = x^n$
RBF basis $\psi_n = \exp{(-\frac{||x-z_n||_2^2}{2l^2})}$ for $l = .02$ and $z_n = -1 + \frac{2n}{N}$
Cosine and Sine $\psi_n = \cos(n\pi x) + \sin(n\pi x)$

Before the implementation of each dictionary, the Python numpy.linalg.solve function is used to find $(\Psi^T\Psi - \lambda\mathbf{I})^{-1}\Psi^T$ for the optimal coefficient vector $\hat{\mathbf{c}}$. In addition, each dictionary and Levy function will be examined with 500 grid points to ensure the curves are smooth.

3. Results

To examine the approximating Cosine, Monomial, and RBF dictionaries each along the Levy curve and the observation points, each dictionary is graphed for a fixed values $N = 10,\ 50,\ \text{and}100$ with 3 varying λ values for each plot.

The Cosine Series dictionary is shown in 1. The left subplot shows the cosine series at a fixed $N = 10$, meaning there are 10 combinations of the Cosine functions with different argument; each curve filled in solid line curve representing distinct λ values. Here, all 3 approximations show similar oscillation patterns, and curvature with the Levy function. In addition, the distance between each curve is close to the Levy function and there are no outliers. However, as N increases to 50 and 100, the Cosine approximation for $\lambda = 1$ overlaps the other two λ, showing rapid oscillation and growing amplitudes as the x points deviate from $x = 0$. In addition, the Cosine basis for $\lambda = 1$ shows overfitting for $N = 100$ as the curve is interpolating the observation point rather than approximating the Levy Function. Thus, the Cosine Series at $N = 10$ show better approximation compared to the other N values and the bigger λ value overlaps the other two, showing overfitting as N increases.

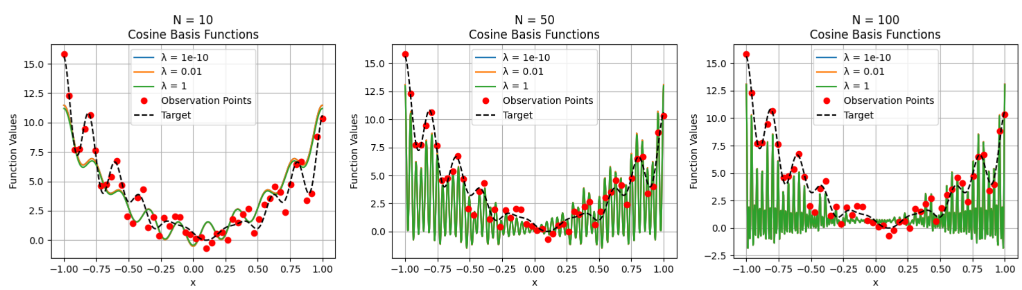

Figure 1. Cosine Series Dictionary: Each plot shows the Cosine Basis for a fixed $N = [10, 50, 100]$ for various $\lambda = [10^{-10}, 10^{-2}, 1]$.

The Monomial dictionary is shown in 2. Again, left subplot shows the Monomial dictionary as a combinations of 10 Monomial basis and each solid line curve representing distinct λ values. Here, each solid curve are distance wise close to the Levy function. However, the Monomial Basis Functions do not oscillate given that the function is not periodic; the curves do not exhibit behaviors of overfitting to observation points for $N = 10$. For both larger N values, the values at the endpoint blows up to overfit for $\lambda = 10^-10$, the smallest λ value, this is called the Runge Phenomena. Specially, at $N = 100$, the curve with $\lambda = 10^-10$ blows up to approximately 1800 while the upper bound for the Levy function on the interval $[-1, 1]$ is 17.5. This shows that as N increases and λ decreases, the Monomial Basis Curves exhibit more contrasting behavior with the Levy function at the endpoints.

The RBF Basis dictionary is shown in 3. Contrary to the other two dictionaries, the Monomial Dictionary exhibits stable behaviors as changes in N and λ values do not change the behaviors of each curve. Notice that the Monomial approximation at $\lambda = 10^-10$ shows a better approximation to the Levy function at the endpoint as it reaches the endpoints, whereas for $N = 10$, all three curves of different λ values does not reach the endpoints of the Levy function, rather they have smaller values. In addition, the overfitting problem seen in the other two dictionaries is not seen

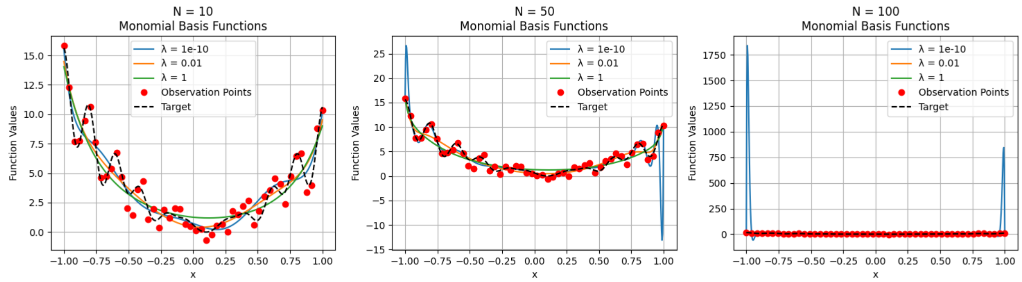

FIGURE 2. Monomial Dictionary: Each plot shows the Monomial Basis for a fixed $N = [10, 50, 100]$ for various $\lambda = [10^{-10}, 10^{-2}, 1]$.

here as there are no rapid oscillation as a byproduct of overfitting.

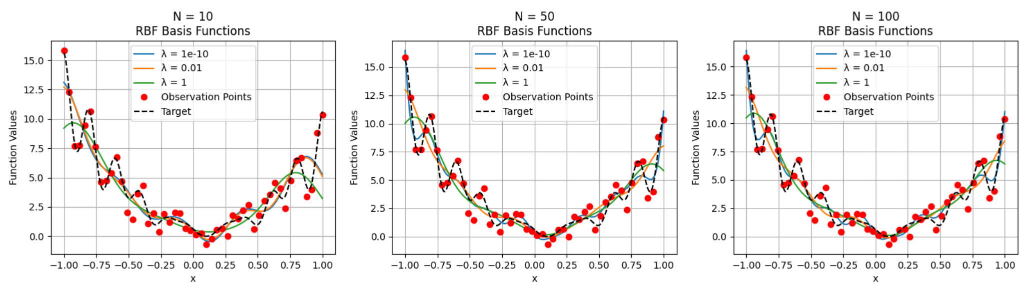

FIGURE 3. RBF Series Dictionary: Each plot shows the RBF Basis for a fixed $N = [10, 50, 100]$ for various $\lambda = [10^{-10}, 10^{-2}, 1]$.

The Cosine and Sine Basis dictionary is shown in 4. Here, $\lambda = 10^{-12},\ 1.2$ and $N = 5, 25$ are plotted for this dictionary. For $N = 5,\ 10,\ 25$, the Cosine and Sine basis functions are well behaved as they do not show signs of overfitting and follow the shape and oscillation of the Levy function. However, for $N = 50$, the approximation curves show signs of rapid oscillation, larger amplitude, while still following the curvature of the Levy Function. This is consistent to the Cosine Series and Monomial dictionaries as larger N values do not approximate as well as smaller N values. For the lambda values, each curve overlaps each other for all values of N.

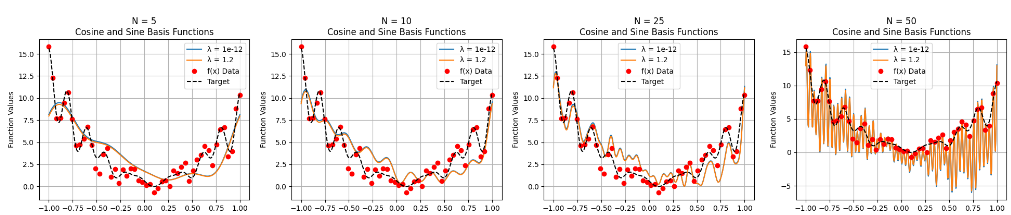

FIGURE 4. Cosine and Sine Dictionary: Each plot shows the Cosine and Sine Basis for a fixed $N = [5, 10, 25, 50]$ for various $\lambda = [10^{-12}, 1, 1.2]$.

4. Summary and Conclusions

In this report, the Cosine Series, Monomial, Radial Basis Function (RBF), and the Cosine and Sine dictionaries were used to approximate the Levy function with the respective observation points. From the results, the RBF dictionary is the most well behaved and follow a similar trend to the Levy function; the range of the regularization value and number of dictionary size do not have large impact on the RBF dictionary. However, the Cosine Series has rapid oscillation and the Monomial dictionary blows up as N increases (the Runge Phenomenon); for larger dictionary sizes, the Monomial curves varies more for different regularization parameter. The Cosine and Sine dictionary is consistent with Monomial and the Cosine Series as they are well behaved and accurate for smaller dictionary sizes. This leads to the question of what other basis function is adequate to approximate the Levy function.

Acknowledgements

I would like to thank Emily Zhang, Tomas Sevcek, and Johan Azambou for helpful discussions.