Project Member Application
ASCENT
2026-27

For
Mathematics Club

Centre for Innovation
IIT Madras

Instructions

General Instructions

- Mention the following details at the start of your application.

Name:		Insti Name:
Roll:	Room:	Hostel:
Phone (WA):	CGPA:	Email:

- Join the Whatsapp group for further updates.
- The recommended font is a standard font size 11-13.
- The applications have to be submitted in PDF format, named as:
  <First_name>_<Roll_Number>_Mathematics_Club_Project_Member.pdf

  For example, Taarun_EE24B069_Mathematics_Club_Project_Member.pdf.
- You can upload the finished applications here.
- You may submit the completed application on or before **11:59 PM, 18 June**

Note:

- Clearly explain your reasoning for each question. Your understanding and approach matter more than the final answer.
- Feel free to use online resources for reading and reference while attempting the application. However, your submitted responses should reflect your own understanding and should not be AI-generated.
- We will consider the use of LLMs to understand concepts and for the R codes as long as you can explain the code. However, other AI generated responses will be penalised . To get started with R please use this.
- It is fine if you don't solve the entire app. Complete as much as you can.
- Focus on the non-bonus questions before attempting the bonus questions.
- If you have any queries, you can reach out to the PLs anytime you want:
  - Anirudh N (CH24B035) : +91 99169 72835
  - Taarun A (EE24B069) : +91 73584 20650

Contents

§1 HR

- Tell us a bit about yourself. Why do you want to join this project? What skills or qualities do you possess that make you suitable for the work involved ?
- Mention all PoRs/activities you are planning to take this year. Weekly, how much time do you think you will be able to commit to this project? How much time will you commit to other PoRs and academics?
- What would keep you motivated throughout the period of two semesters? How would you ensure that you are consistent in contributing to the project throughout the tenure?

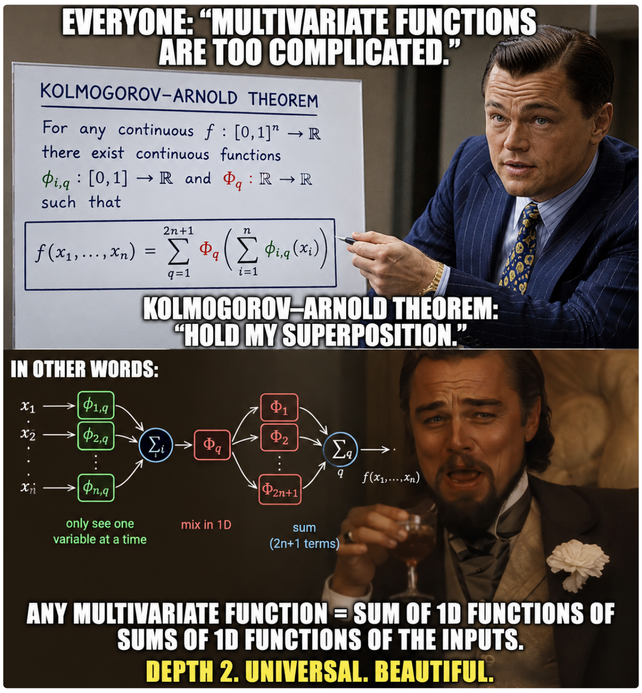
§2 Kolmogorov - Arnold Theorem

Kolmogorov - Arnold Theorem states that any multivariate continuous function in a bounded domain can be represented as a superposition of continuous functions of one variable and addition.

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{2n+1} \phi_i \left(\sum_{j=1}^{n} \psi_{ij}(x_j) \right)$$

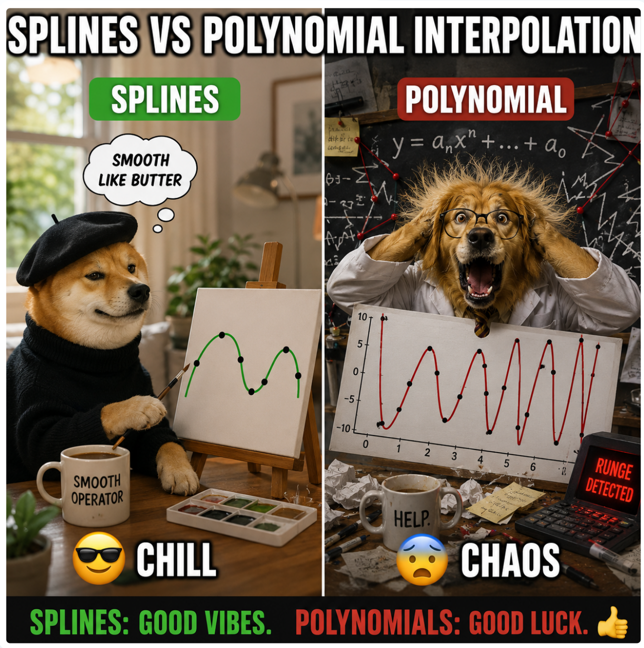
But unfortunately, the theorem only guarantees continuity, but the 1-D functions might be highly non-smooth. Still, some cases might admit elegant decompositions. For the following functions, find suitable functions ϕ_i and ψ_{ij} satisfying the above representation.

1. $\dfrac{x_1^2}{x_2^3}, \quad x_2 \neq 0$
2. $\sin(2x_1)\cos(x_1^2 + x_2^2)$
3. $\log_{x_1}(x_2), \quad 3 \leq x_1 \leq 4, 1 \leq x_2 \leq 2$
4. $u = \dfrac{x_1 + x_2}{1 - x_1 x_2}, v = \dfrac{x_1 - x_2}{1 + x_1 x_2}, z = \dfrac{u}{\sqrt{1+u^2}} + \dfrac{1}{\sqrt{1+v^2}}, \quad 0 < x_1 < 1, 0 < x_2 < 1$



§3 Splines!!!

A spline is a function defined piecewise by polynomials. The value of independent variable at the places where the polynomial pieces meet are called knots.

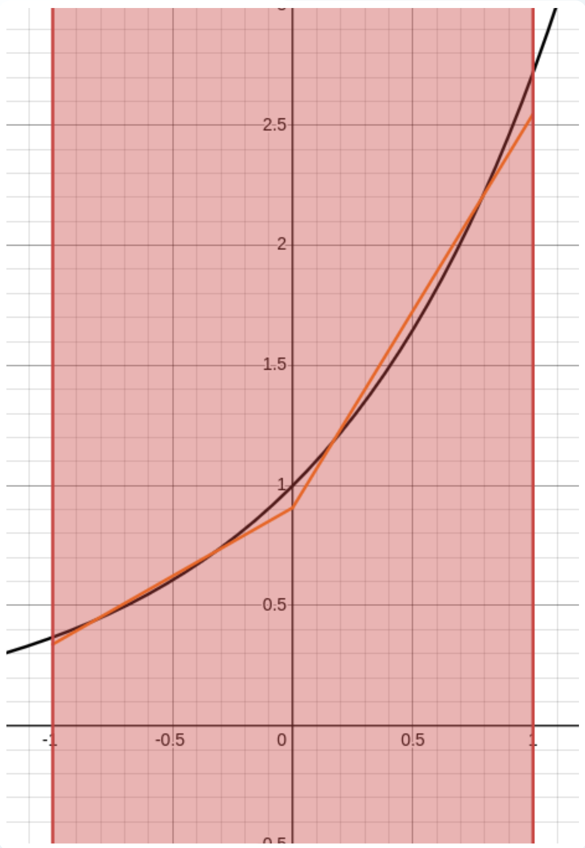


1. a) Mr.K wishes to create a 2nd degree spline (All pieces are quadratics) that is C^1 continuous $\forall x$. He also wants the spline to have knots (points where the polynomials meet) only at $x = 0$, $x = 1$, $x = 2$ and $x = k$. He decides to fix the function values at knots to be $y = 0$ at $x = 0$, $y = 1$ at $x = 1$, $y = 3$ at $x = 2$ and $y = 0$ at $x = k$. He takes $f(x) = 0 \; \forall \; x \le 0$ and $f(x) = 0 \; \forall \; x \ge k$ Find the value of k. Does the obtained value of k make sense?

   b) He decides to introduce one more knot at $x = 4$ and fixes the function value at $x = 4$ to be 3. Having the same conditions as before, find the value of k?

   c) If instead of fixing the function value at $x = 4$ to be 3, He fixes it to be a hidden number z. Having the same conditions as before, what is the range of z to yield a proper C^1 continuous spline.

2. a) Find the linear spline in the interval $x \in [-1, 1]$ that is C^0 continuous $\forall x \in (-1, 1)$, has knots only at $x = -1$, $x = 0$ and $x = 1$ and minimizes the squared error from e^x. $(min\left(\int_{-1}^{1}(f(x) - e^x)^2 dx\right))$



b) Instead of fixing knot at $x = 0$, if we were allowed to choose the knot location, how does the quantity being minimized change? Write its expression in terms of y_1(y-value at $x = 1$), y_{-1}(y value at $x = -1$), k(the new knot is at $x = k$) and y_k(y-value at $x = k$). You need not find the spline that minimizes the squared error for this part.

3. a) Prove that if we have a set of 1-D splines having k different knots in total, on applying a affine transformation on them (multiplying each spline by some constant, adding them up and then adding a constant), followed by ReLU(x), we could generate $k + 1$ new unique knots.

   b) Let:

$$f_1(x) = -ReLU(-x + 1) - ReLU(x - 1) \qquad \text{(Has 1 knot at } x = 1\text{)}$$
$$f_2(x) = ReLU(x - 2) \qquad \text{(Has 1 knot at } x = 2\text{)}$$
$$f_3(x) = ReLU(x - 3) \qquad \text{(Has 1 knot at } x = 3\text{)}$$

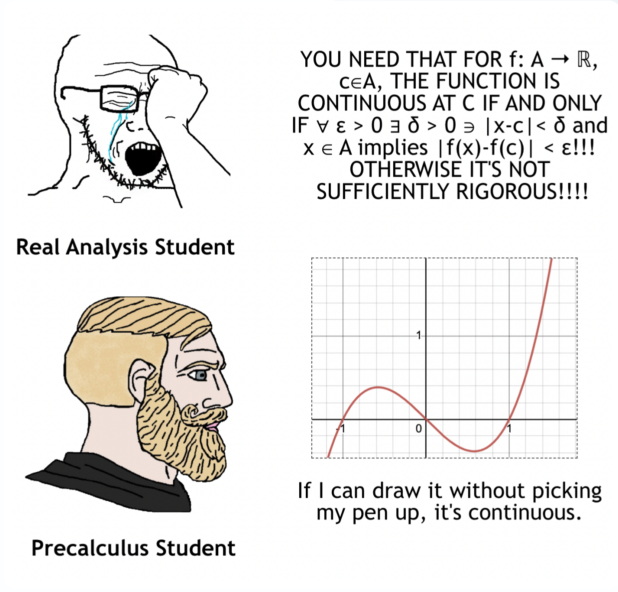
   Find an affine transformation following which on applying ReLU(x), we get 4 new knots.

   c) Assume an MLP (Multi-Layer Perceptron) with ReLU activation function everywhere with a one dimensional input, x. Let the total number of knots generated upto the k-th layer be m_k. Prove that:

$$m_l \leq m_{l-1} + n_l(m_{l-1} + 1)$$

   $\forall l \in \mathbb{W}$ where n_l is the number of nuerons in the l^{th}- layer. The $l = 1$ layer directly works on input, $m_0 = 0$.

4. Basis function for B-splines can be constructed using the Cox-de Boor recursion formula, You could read about it here.



a) Show that if multiplicity (number of times a knot is repeated in the knot vector) of all knots is 1, then the basis functions generated by the Cox-de Boor recursion formula have C^{p-1} continuity $\forall x$ where p is the degree of the spline.

b) Let a degree p B-spline have a knot t_k with multiplicity m, prove that continuity at t_k becomes C^{p-m}.

5. a) Read up on De-Boor's Algorithm, how does it speed up computation when compared with using Cox de-Boor recursion formula explicitly?

§4 Hypothesis Testing: Are We Sure About Anything?

Hypothesis testing is a statistically supported method to gauge the correctness of your claims. Given a dataset, we want to decide between two competing statements about an unknown population parameter θ:

- H_0: $\theta = \theta_0$.
- H_1: $\theta \neq \theta_0$, or $\theta > \theta_0$.

We collect data, compute a *test statistic* T, and probabilistically quantify if H_0 were true, how surprising would a value of T be?

§ The Null and the Restless

Say you have a sample $y_1, \ldots, y_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ with σ^2 known, and you want to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$.

The test statistic is:

$$Z = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}}$$

1. Explain hypothesis testing. What are the kinds of hypotheses that can be tested? Explain the significance and power of the tests. How are they different for the nature of hypotheses mentioned above?

2. No decision rule provides complete accuracy under all environments.
   (a) For the one-sample Z-test above, derive an expression for the power as a function of the true mean μ, the hypothesised mean μ_0, σ, n, and α.
   (b) Sketch (or describe) how the power curve changes as (i) n increases, (ii) $|\mu - \mu_0|$ increases, and (iii) α increases.
   (c) Why is it generally not possible to simultaneously minimise both error types without increasing the sample size?

3. When σ^2 is unknown, how do we quantify the test statistic?
   (a) What distribution does it follow? How does it depend on the number of data points?
   (b) A study measures the resting heart rate (bpm) of 20 individuals after a mindfulness programme. The sample mean is $\bar{y} = 68.4$ bpm and the sample standard deviation is $s = 7.2$ bpm. Test whether the true mean differs from 72 bpm at the 5% significance level. State H_0, H_1, compute t, find the p-value (you may express it in terms of the CDF), and state your conclusion.

4. Suppose two independent groups are measured. Group A (sample size n_A, mean $\bar{y}_A$, variance s_A^2) and Group B (n_B, $\bar{y}_B$, s_B^2). We test $H_0 : \mu_A = \mu_B$.
   (a) Write down the pooled two-sample statistic under the assumption of equal variances.
   (b) What is Welch's t-test and when should it be preferred over the pooled version?
   (c) What would you use if the two groups are not independent?

5. Show that for the one-sample case, the $(1 - \alpha) \times 100\%$ confidence interval for μ is exactly the set of values μ_0 that would *not* be rejected at level α in the two-sided t-test. This duality is fundamental.

6. **Bonus:** Explain the Generalised Likelihood Ratio Test.

§5 Predicting The Future

So far, we have assumed that our observations are independent and identically distributed. But what if the data has a natural ordering in time, and consecutive observations are correlated?

A time series is a sequence of observations $\{Y_t\}_{t=1}^{T}$ indexed by time. Examples: stock prices, temperature readings. The statistical challenge is to model the temporal dependence in the data so that we can describe, explain, and forecast. **For questions involving statistical tests, explain the statistic used and the distribution that follows.**

§ Is Your Data Even Stationary?

Before fitting any model, we must understand whether the statistical properties of the series change over time. A time series $\{Y_t\}$ is (weakly) stationary if:

1. $\mathbb{E}[Y_t] = \mu$ for all t (constant mean),
2. $\text{Var}(Y_t) = \sigma^2 < \infty$ for all t (constant variance),
3. $\text{Cov}(Y_t, Y_{t+h}) = \gamma(h)$ depends only on the lag h, not on t.

1. (a) Why is assessing stationarity before fitting time series models important?
   (b) What is differencing, and how does it help transform a non-stationary series into a stationary one? Apply this to a vanilla random walk and show that it is stationary.
2. Additive decomposition is as follows:
   $$Y_t = T_t + S_t + R_t$$
   where T_t is the trend, S_t is the seasonal component, and R_t is the residual.
   (a) Explain why a multiplicative decomposition $Y_t = T_t \cdot S_t \cdot R_t$ might be preferred when the seasonal fluctuations grow with the level of the series.
   (b) Using R's `decompose()` or `stl()` function (you may use the `AirPassengers` dataset), decompose the series and plot the four components. Comment on what you observe.
3. (a) Explain the difference between the ACF and the PACF. How can they be used for model identification ?
   (b) How are they estimated and what distributions do they follow (Assuming Gaussian noise)?
   (c) Generate 500 observations of an AR(2) process $Y_t = 0.6Y_{t-1} - 0.3Y_{t-2} + \varepsilon_t$ in R (`arima.sim`). Plot the ACF and PACF. Do the empirical plots match theory?
4. The Augmented Dickey-Fuller test:
   H_0 : the series has a unit root
   H_1 : the series is stationary.
   (a) State the regression equation for the ADF test and explain each term.
   (b) Apply the ADF test (R package `tseries: adf.test`) to the `AirPassengers` series before and after log-transformation and differencing. At what stage does the series become stationary?

§ ARIMA

An ARMA(p, q) model for a stationary series:

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

where $\varepsilon_t \overset{iid}{\sim} (0, \sigma^2)$.

5. (a) What is an ARMA model? Explain its terms.
   (b) Explain the AIC and BIC formulae for a fitted ARIMA model. What do the terms penalise, and why does BIC tend to select sparser models than AIC?
   (c) Given ACF and PACF plots from a stationary series, describe the procedure you would follow to select initial values of p and q.
   (d) What does R's `auto.arima()` function do? Is this always a reliable substitute for manual identification?
6. Using the `AirPassengers` dataset (apply `log()` and then difference once to achieve stationarity):
   (a) Fit an ARIMA model using `auto.arima()` from the `forecast` package. Report the selected (p, d, q) order and the AIC.
   (b) Perform residual diagnostics: plot the residual ACF, run a Ljung–Box test (R: `Box.test`), and check whether residuals look like white noise. Why does white noise in residuals matter?
   (c) Plot forecasts for the next 24 months with 80% and 95% prediction intervals. Comment on the intervals.

§6 Resources

- Introduction to Hypothesis testing.
- Introduction to Information Criteria.
- The first few lectures are useful for understanding Time Series Analysis.
- An introductory blog on B-splines
- Wiki B-splines
- A youtube video explaining the Kolmogorov-Arnold Theorem