

I, _____, pledge that this exam is completely my own work, and that I did not take, borrow or steal any portions from any other person. I understand that if I violate this honesty pledge, I am subject to disciplinary action pursuant to the appropriate sections of the San Diego State University Policies.

Signature

Be sure to show all your work or include a copy of your programs.

1. This problem examines a spline clamped in the middle of the data. Consider the data given in the table below.

x	y	x	y	x	y
0	0	7	3.7	14	4.5
1	5.2	8	4.8	15	3.1
2	2.3	9	5.5	16	5.7
3	6.2	10	6	17	3.2
4	7.3	11	8.7	18	4.4
5	5.1	12	5.2	19	2.1
6	4.5	13	3.1	20	0

a. Create a natural spline through the data. Find the y -values at $x = 2.5, 6.5, 11.5$, and 18.5 . If $S_n(x)$ is this natural spline, $S_n(x)$, then use the coefficients that you found to exactly evaluate

$$\int_3^{15} S_n(x) dx.$$

b. We again allow the curve to have natural boundary conditions (no force) on the endpoints, $x = 0$ and 20 . However, we clamp this spline in the middle with the curve forced to be flat. Thus, $y(10) = 6$ and $y'(10) = 0$. Find this spline, $S_f(x)$, and create a graph showing the data, the natural spline through all points, and $S_f(x)$ on the domain of the data. Use $S_f(x)$ to find the y -values at $x = 2.5, 6.5, 11.5$, and 18.5 . If $S_f(x)$ is this spline, then use the coefficients that you found to exactly evaluate

$$\int_3^{15} S_f(x) dx.$$

c. Use the Composite Simpson's Rule with $h = 1$ to evaluate the integral from $x = 3$ to $x = 15$, passing through the data. Also, use the spline $S_f(x)$ to generate the midpoints between each of the integers from 3 to 15, then use the CSR to approximate

$$\int_3^{15} S_f(x) dx.$$

Compare all 4 integral values you obtain.

2. a. The pathogen, *Staphylococcus aureus*, is linked to many cases of food poisoning. Data were collected on a growing culture over 5 hr here at SDSU.

t (hr)	0	0.5	1	1.5	2	2.5
OD ₆₅₀	0.032	0.039	0.069	0.110	0.170	0.229
t (hr)	3	3.5	4	4.5	5	
OD ₆₅₀	0.261	0.288	0.309	0.327	0.347	

The population (measured in OD₆₅₀) best fit a *logistic growth model* given by:

$$S(t) = \frac{0.3450}{1 + 12.85e^{-1.24t}}, \quad t \in [0, 5].$$

Create a graph of the data and $S(t)$. Use your techniques from Calculus to compute $S'(2)$ and $S'(3)$.

b. We want to estimate $S'(2)$ and $S'(3)$ with our **3 point** and **5 point** differentiation formulas. Let $h = 0.25$ with the function, $S(t)$ above. Specifically, for the **3 point** differentiation formula use the **forward** and **central** formulas to estimate $S'(2)$ and $S'(3)$ and find the absolute error between these estimates and the actual values found in Part a. Also, for the **5 point** differentiation formula use the second **forward**:

$$S'(x_0) \approx \frac{1}{12h} (-3S(x_0 - h) - 10S(x_0) + 18S(x_0 + h) - 6S(x_0 + 2h) + S(x_0 + 3h)),$$

and **central** formulas to estimate $S'(2)$ and $S'(3)$ and find the absolute error between these estimates and the actual values found in Part a.

c. For this part of the problem, use the **data from the table** given above along with the **3 point** differentiation formulas to estimate the growth, $S'(t)$, of *S. aureus*. Except for $t = 0$ and $t = 5$, use the data in the **3 point** central differentiation formula to estimate the derivative of $S(t)$ (with $h = 0.5$). At $t = 0$ and $t = 5$, use the forward and backward **3 point** differentiation formulas, respectively (with $h = 0.5$). Find the absolute error of these data based approximations to the derivative at $t = 2$ and 3 to the actual derivatives, $S'(2)$ and $S'(3)$, from the function in Part a.

d. Find the best fitting **cubic function**, $C(t)$, through the approximations that you found in Part c. Graph the estimates of the derivatives from Part c as points ($t \in [0, 5]$), then overlay $S'(t)$ (coming from the function $S(t)$ in Part a) and $C(t)$ from this part. Determine $C'(t)$ and find the maximum of this cubic (using Calculus). Also, find the absolute difference in the times and growths computed with $C(t)$ and the point of inflection for $S(t)$ from Part a. (Note: The maximum of $C(t)$ is the maximum growth rate from this approximation, and this should correlate to the point of inflection for $S(t)$, so compute $|t_a - t_b|$, where $C'(t_a) = 0$ and $S''(t_b) = 0$, and find $|C(t_a) - S'(t_b)|$.)

3. Consider the following table of data.

x	y	x	y	x	y
0	23.22	3.5	18.07	7	13.04
0.5	23.08	4	18.47	7.5	11.87
1	22.74	4.5	16.66	8	9.57
1.5	22.54	5	15.91	8.5	8.75
2	21.16	5.5	13.50	9	6.63
2.5	19.12	6	13.17	9.5	5.80
3	19.70	6.5	12.49	10	6.51

If $x = [x_0, \dots, x_n]^T$, then create the matrices A_m for $m = 1, 2, 3, 4$, with

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^m \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^m \end{pmatrix}.$$

(**Don't write** these large matrices in your report.) These matrices are central to the **normal equations** for the $(m-1)^{st}$ degree polynomial and are given by

$$A_m^T A_m a = A_m^T y,$$

where $a = [a_0, \dots, a_{m-1}]^T$ and from the data, $y = [y_0, \dots, y_n]^T$.

a. We want to find the least squares best linear fit to these data. Compute and record the matrix $A_1^T A_1$ in your report. Solve the normal equations, giving the coefficients for the best linear model. Write clearly the equation for the best linear model fitting these data. Also write your sum of square errors for the linear fit. Give all answers in decimal form with at least 5 significant figures.

b. Repeat process in Part a to find the least squares best fit of quadratic, cubic, and quartic polynomials. Write the matrices

$$A_m^T A_m, \quad \text{for } m = 2, 3, 4.$$

Solve the associated normal equations. Write clearly the equations for the best quadratic, cubic, and quartic model fits to the data. Also, give the sum of square errors for each of these models. Present a graph showing the data and all of the models computed above.

c. Use the *Bayesian Information Criterion (BIC)* and the *Akaike Information Criterion (AIC)* to determine the best models. Write the BIC and AIC numbers associated with the linear, quadratic, cubic, and quartic polynomials, then use these numbers to state which model is best (so should be used) according to each of these criterion.

4. a. Below is a Table of Metabolism (kcal) and Weight (kg) for Various Animals

Animal	Weight	Metabolism	Animal	Weight	Metabolism
Mouse	0.021	3.6	Dog	24.8	875
Rat	0.282	28.1	Dog	23.6	872
Guinea pig	0.41	35.1	Goat	36	800
Rabbit	1.52	83	Sheep	46.4	1254
Rabbit	2.46	119	Sheep	46.8	1330
Rabbit	3.57	154	Chimpanzee	38	1090
Rabbit	4.33	191	Woman	54.8	1224
Rabbit	5.33	233	Woman	57.9	1320
Cat	3	152	Cow	300	4221
Macque	4.2	207	Cow	435	8166
Dog	6.6	288	Heifer	482	7754
Dog	14.1	534	Cow	600	7877

a. Use our techniques from class to find the best linear, quadratic, and cubic models in the least squares sense. Write the models including their coefficients and include the least sum of square errors for each of the models. Give all answers in decimal form with at least 5 significant figures.

b. Max Kleiber (1947)¹ wrote an article about the relationship between body weight and metabolism of animals. He argued that there was an allometric or power law relationship of the form

$$M = aW^k,$$

which can be written in the logarithmic form

$$\ln(M) = \ln(a) + k \ln(W).$$

This a linear model in the logarithms of the data. Find the best linear least squares fit to the logarithms of the data and determine the constants, a and k . Compute the the sum of square errors

$$\sum_{i=0}^N \left(M_i - aW_i^k \right)^2. \quad (1)$$

Compare this model to your best fitting models in Part a. Create graphs of the polynomial models and this allometric model with the data. On one graph use the standard linear scale, then use MatLab's `loglog` graph to show how these models appear with logarithmic scales.

c. Use MatLab's `fminsearch` to find the nonlinear least squares best fit with Eqn. (1). Give the best fitting constants, a and k and compute the the sum of square errors. Compare this result to the other models in Parts a and b. Create a graph comparing the models in Parts b and c, using MatLab's `loglog` graph style.

d. Use the graphs above and the SSEs to compare and contrast all of the models. Write a short paragraph discussing the strengths and weaknesses of the models. (Address the issue of what the models predict as $w \rightarrow 0$.) What do the allometric models say about the relation between weight and metabolism? What does this say about the proportional heat production of small animals compared to large animals? Can you explain the coefficient k ? (Detailed explanations are **not** expected.)

¹Max Kleiber (1947), "Body size and metabolic rate," *Physiological Reviews*, **24**, 511-541

5. Remifentanil is an opioid drug that acts very quickly and is rapidly degraded by esterases. A bolus injection of 15 μ /kg was administered to a patient and the plasma concentration was measured over a period of time. Data from the experiment is given in the table below.

t	R	t	R	t	R
1	53.15	7	5.56	25	0.76
2	28.25	8	3.96	30	0.57
3	18.82	10	3.02	40	0.39
4	11.19	12	2.2	60	0.23
5	8.93	15	1.72	90	0.13
6	7.13	20	1.17		

where t is in min and R is the concentration of remifentanil in ng/ml.

a. Perform a least squares fit of a quadratic (3-parameter) model to the remifentanil concentration in the plasma,

$$Q(t) = q_0 + q_1t + q_2t^2.$$

Write the equation for this best-fitting model. Also, consider a 3-compartment model for disappearance of remifentanil given by

$$C(t) = a_1e^{-\lambda_1t} + a_2e^{-\lambda_2t} + a_3e^{-\lambda_3t}.$$

Suppose that other measurements suggest that the *rapid distribution phase* has $\lambda_1 = 0.6$, the *middle distribution phase* has $\lambda_2 = 0.15$, and the *terminal/elimination phase* has $\lambda_3 = 0.025$. Perform a least squares fit to find the 3-parameters, a_1 , a_2 , and a_3 , and write the formula with the best-fitting parameters in your model. Write the least sum of square errors for each model. Create a graph of the data and these two models using a logarithmic scale for the vertical axis (semilog graph). Which model appears to fit the data best?

b. Perform the **exponential peeling** procedure separating out the data over the ranges $t \in [0, 7]$, $t \in [8, 25]$, and $t \in [30, 90]$. Give the best fitting exponential models (using a linear least squares fit to the logarithms of the concentrations) over these ranges, where you start with the *terminal/elimination phase* and subtract each model from the data before computing the next exponential fit. Give the complete 3-compartment model using this exponential peeling procedure. Include the sum of square errors for this model.

c. Start with the 6 parameters from the exponential peeling (3 coefficients and 3 exponents) and use MatLab's `fminsearch` (Nonlinear Least Squares fit) to find the best fitting parameters for the 3-compartment model. This procedure fails with a direct least squares fit of the model to the data, so perform a least squares best fit of the logarithm of the model to the logarithm of the data. Write this best fitting model with its parameters and give the sum of square errors. Write a brief paragraph comparing and contrasting the models in the entire problem and giving your opinion on which ones are best from both how they fit the data and the computational effort. Create a semilog graph including the data, the 3-compartment model from Part a, the exponential peeling model from Part b, and the Nonlinear least squares best fit model.