

Imagine that you have been hired by a city to create a model for violent crime in their city. The city has given you three potential sampling schemes in `param-uncertainty-sampling.txt`. Given the model:

$$f(t_i) = \theta_1 \sin\left(\frac{2\pi}{365}t_i\right) + \theta_2 \sin\left(\frac{2\pi}{30}t_i\right) + \theta_3 \sin\left(\frac{2\pi}{7}t_i\right) + \theta_4 t_i + \theta_5$$

- (a) Construct the system matrix A for each set of points
- (b) Calculate the Fisher Information Matrix using the equation $A^T A$ for each set of points (Note that we have not yet generated a fit, so we are setting σ^{-2} to 1 for simplicity's sake).
- (c) Invert the Fisher Information Matrix to yield the covariance matrix. Report the diagonal values for each covariance matrix, which are the uncertainties associated with each θ value.
- (d) If the city wants to constrain θ_4 , which sampling scheme should they choose and why?
- (e) In regression problems, σ^2 is calculated by scaling the sum of the square errors by the degrees of freedom (the number of data points minus the number of parameters in your model). Under what circumstances would your answer to part d change after calculating σ^2 ?