Stanford AA 203: Optimal and Learning-based Control Problem set 0

Problem 1: Poisson Maximum Likelihood

Suppose we observe the number of customers to a store over n days $x_1, x_2, ..., x_n$, and we want to fit a Poisson distribution to this data. The Poisson distribution is a distribution over non-negative integers with a single parameter $\lambda \geq 0$. It is often used to model arrival times of random events or count the number of random arrivals within a given amount of time. It has probability mass function:

$$
\mathbb{P}_{\lambda}[X = k] = \frac{e^{-\lambda}\lambda^k}{k!} \text{ when } X \sim \text{Poi}(\lambda).
$$

One way to do this is via Maximum Likelihood, where we choose the parameter of the Poisson distribution to maximize the probability that the data $x_1, x_2, ..., x_n$ appears. This can be done by maximizing the log-likelihood of the dataset $x_1, ..., x_n$ with respect to λ . The log-likelihood of $x_1, ..., x_n$ under the Poisson model is

$$
f_{\lambda}(x_1,...x_n) := \sum_{i=1}^n \ln\left(\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}\right).
$$

Compute the Maximum Likelihood estimator $\hat{\lambda}$ by finding a solution to

$$
\arg\max_{\lambda\geq 0} f_{\lambda}(x_1,...,x_n).
$$

Problem 2: Discrete Linear Systems

Consider the discrete linear system $x_{t+1} = Ax_t + Bu_t$, where

$$
A = \begin{bmatrix} \frac{4}{5} & 0 & 0 \\ 0 & \sqrt{3} & 1 \\ 0 & -1 & \sqrt{3} \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}.
$$

- a) In the absence of control (i.e. $u_t = 0$), is this system stable? Why or why not?
- b) Design a linear feedback controller $u_t = Kx_t$ for some fixed matrix $K \in \mathbb{R}^{2 \times 3}$ so that the closed loop system will be stable.

Problem 3: Linear Regression

Recall that the least squares solution to $\min_x ||Ax - b||_2^2$ $\frac{2}{2}$ is given by the normal equation $x^* = (A^{\top}A)^{-1}A^{\top}b$.

a) Suppose in addition to finding an x so that Ax is close to b, we prefer x to be "small" as measured by $x^{\top} \Lambda x$, where Λ is a positive definite matrix. This gives rise to the ridge regression problem:

$$
\min_x \|Ax - b\|_2^2 + x^\top \Lambda x.
$$

Derive the normal equation (i.e. closed form solution) for the ridge regression problem.

b) We obtain measurements $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ on n asteroids, where x_i, y_i are estimates of the diameter and mass of the ith asteroid respectively. If the asteroids were radially symmetric and uniformly dense, then by a volume argument, we could deduce that $y_i = \frac{4\pi}{3}$ $rac{\pi}{3}$ $\left(\frac{x_i}{2}\right)$ $\left(\frac{v_i}{2}\right)^3$. The asteroids however, are not radially symmetric nor uniformly dense, but we still suspect that x, y exhibit a cubic relationship, i.e. $y = p(x)$ where p is a cubic polynomial. Using the data $\{(x_i, y_i)\}_{i=1}^n$ in prob3data.csv, find the coefficients c_0, c_1, c_2, c_3 so that $p(x) := c_0 + c_1 x + c_2 x^2 + c_3 x^3$ is the least squares cubic estimator of y from x.

Problem 4: Gradient Methods

Recall the polynomial fitting approach from Problem 3. Suppose we want a solution that is robust to outliers. One way to do this is to replace the ℓ_2 norm in least squares with an ℓ_1 norm, where for a vector $x \in \mathbb{R}^n$, its ℓ_1 norm is given by $||x||_1 := \sum_{i=1}^n |x_i|$. This gives rise to the following optimization problem:

$$
\min_{x} \|Ax - b\|_1. \tag{1}
$$

One common technique for optimization is called Gradient Descent which uses the function's derivative to iteratively reduce the objective value. Given a function f and an initial starting point x_0 , Gradient descent produces a sequence of iterates $x_1, x_2, ...$ until convergence according to the following rule:

$$
x_{k+1} = x_k - \alpha \nabla f(x)
$$

where $\alpha \geq 0$ is the step size. Using the same A, b from Problem 3, implement a Gradient Descent algorithm to solve [\(1\)](#page-1-0).

Learning goals for this problem set:

- Problem 1: To review unconstrained convex optimization.
- Problem 2: To review stability analysis of discrete linear systems.
- Problem 3: To review linear regression techniques and applications
- Problem 4: To review first order optimization methods.