Please submit your homework with codes (hard copy) in class and upload the corresponding codes to the Blackboard. Problems marked with * will be graded in detail and they are worth 50% of the total score. Remaining problems, worth the remaining 50% of the total score, will be given full mark if reasonable amount of work is shown.

For this homework, use R for programming parts unless otherwise specified.

1. * The Cauchy$(\theta, 1)$ has density
$$p(x-\theta) = \frac{1}{\pi\{1+(x-\theta)^2\}}.$$
   (a) If $x_1, \ldots, x_n$ form an i.i.d. sample, show that
$$\begin{aligned} l(\theta) &= -n\log\pi - \sum_{i=1}^{n}\log\{1+(\theta-x_i)^2\}, \\ l'(\theta) &= -2\sum_{i=1}^{n}\frac{\theta-x_i}{1+(\theta-x_i)^2}, \\ l''(\theta) &= -2\sum_{i=1}^{n}\frac{1-(\theta-x_i)^2}{\{1+(\theta-x_i)^2\}^2}. \end{aligned}$$
   (b) Show that the Fisher information is $I(\theta) = \frac{n}{2}$.
   (c) Use the following data, graph the log likelihood function: -13.87, -2.53, -2.44, -2.40, -1.75, -1.34, -1.05, -0.23, -0.07, 0.27, 1.77, 2.76, 3.29, 3.47, 3.71, 3.80, 4.24, 4.53, 43.21, 56.75.
   (d) Find the MLE for θ using the Newton-Raphson method (use the same data set as above). Try the following starting points: -11, -1, 0, 1.4, 4.1, 4.8, 7, 8, and 38. Compare your results.
   (e) First use Fisher scoring to find the MLE for θ, then refine your estimate using Newton-Raphson. Try the same starting points as above. Compare your results with the previous ones.

2. Consider the following probability density function:
$$p(x) = \frac{1-\cos(x-\theta)}{2\pi}, \quad 0 \le x \le 2\pi,$$
   where θ is a parameter between $-\pi$ and π. Use this i.i.d. sample to answer the following questions: 0.52, 1.96, 2.22, 2.28, 2.28, 2.46, 2.50, 2.53, 2.54, 2.99, 3.47, 3.53, 3.70, 3.88, 3.91, 4.04, 4.06, 4.82, 4.85, 5.46.
   (a) Graph the log likelihood function.
   (b) Find the method-of-moments estimator for θ. Denote it as $\hat{\theta}_{\text{moment}}$.
   (c) Find the MLE for θ using Newton-Raphson with $\theta_0 = \hat{\theta}_{\text{moment}}$.
   (d) What solutions do you find when you start at $\theta_0 = -2.7$ and $\theta_0 = 2.7$?
   (e) Repeat the above using 200 equally-spaced starting values between $-\pi$ and π. Partition the values into sets of attraction; i.e., divide the set of starting values into separate groups, with each group corresponds to a unique computed MLE.

3. * In chemical kinetics the Michaelis-Menten model is used for modeling the relation between the initial velocity y of an enzymatic reaction and the substrate concentration x. The model is
$$y = \frac{\theta_1 x}{x+\theta_2} + \epsilon, \tag{1}$$
   where θ_1 and θ_2 are model parameters and ϵ are iid zero mean normal errors. The following data set is given:

substrate concentration x (ppm)	velocity y [(counts/min)/min]	
0.02	47	76
0.06	97	107
0.11	123	139
0.22	152	159
0.56	191	201
1.10	200	207

We wish to estimate θ_1 and θ_2 by nonlinear least squares; i.e., $\hat{\theta}_1$ and $\hat{\theta}_2$ are defined as the joint minimizer of

$$\sum_{i=1}^{n} \left(y_i - \frac{\theta_1 x_i}{x_i + \theta_2} \right)^2 . \tag{2}$$

(a) A quick way for finding rough estimates for θ_1 and θ_2 is to invert the relationship in (1) and obtain a simple linear regression setting. That is, ignoring ϵ,

$$\frac{1}{y} = \frac{x + \theta_2}{\theta_1 x} = \frac{1}{\theta_1} + \frac{\theta_2}{\theta_1}\frac{1}{x} \Rightarrow y^* = \beta_0 + \beta_1 u,$$

where $y^* = 1/y, \beta_0 = 1/\theta_1$, $\beta_1 = \theta_2/\theta_1$ and $u = 1/x$. Estimate θ_1 and θ_2 via estimating β_0 and β_1 with least squares.

(b) Implement a Newton-Raphson algorithm for estimating θ_1 and θ_2 via the minimization of (2). Use your answers from (a) as initial estimates.

(c) Repeat (b) with the steepest descent algorithm.

(d) Repeat (b) with the Gauss-Newton algorithm.

4. In this question we want to solve the traveling salesman problem discussed in class using simulated annealing. The matrix below summarizes the distances between any two of 15 cities; e.g., City B and City E are 7 units apart, while City C and City O are 4 units apart.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
A	0	1	2	4	9	8	3	2	1	5	7	1	2	9	3
B	1	0	5	3	7	2	5	1	3	4	6	6	6	1	9
C	2	5	0	6	1	4	7	7	1	6	5	9	1	3	4
D	4	3	6	0	5	2	1	6	5	4	2	1	2	1	3
E	9	7	1	5	0	9	1	1	2	1	3	6	8	2	5
F	8	2	4	2	9	0	3	5	4	7	8	3	1	2	5
G	3	5	7	1	1	3	0	2	6	1	7	9	5	1	4
H	2	1	7	6	1	5	2	0	9	4	2	1	1	7	8
I	1	3	1	5	2	4	6	9	0	3	3	5	1	6	4
J	5	4	6	4	1	7	1	4	3	0	9	1	8	5	2
K	7	6	5	2	3	8	7	2	3	9	0	2	1	8	1
L	1	6	9	1	6	3	9	1	5	1	2	0	5	4	3
M	2	6	1	2	8	1	5	1	1	8	1	5	0	9	6
N	9	1	3	1	2	2	1	7	6	5	8	4	9	0	7
O	3	9	4	3	5	5	4	8	4	2	1	3	6	7	0

Implement a simulated annealing algorithm to find the optimal traveling path. To begin with, set $\alpha(\tau) = p\tau$ with $p = 0.999$, $\beta(m) = 100$ (i.e., $m_j = 100$ for all j), and use a uniform distribution for the proposal density $g_t(|_t)$. Also, set the initial temperature $\tau_1 = 400$. Once your algorithm is working, play around with different values of τ_1, p etc to see how the algorithm behaves. (The shortest path is of length 17.)