ORSUS Research

Mandatory Math Exam

November 2017

Mandatory Exam

Orsus Research Exam
Date: ____________________
Time Limit: 120 Minutes
Start time: ____________________
End time: ____________________

Name (Print): ____________________
University and Year: ____________________

This exam contains 7 pages (including this cover page) and 4 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

The exam is designed to be finished in 120 minutes by an ideal candidate. However, we allow all the examinees to take the exam home and finish it before the due time.

You can use your textbooks, notes, or even search on the Internet when working on this exam. You may even talk to at most two of your friends to finish the exam. However, under each question, you **MUST** write down all the references you have used or the name of people you have talked with to solve it. Any violation of this rule is **NOT** acceptable.

Some of the questions may require some efforts even inspirations. We do not expect or require you to finish all the problems. Please show your work and thoughts on each problem of this exam. If you need more space, use the back of the pages; clearly indicate when you have done this. If you still have questions about the exam, please contact `mbettigole@orsusresearch.com` for help.

Problem	Points	Score
1	15	
2	35	
3	25	
4	25	
Total:	100	

1. (15 points) (a) (5 points) Yao constructed two linear regression models to forecast stock returns. One model is associated with a R^2 of 0.58, while the other has a R^2 of 0.25. Can we argue that the first model is better than the second? Explain your argument very briefly with at most 100 words.

   (b) (10 points) A is a $N \times T$ stock return matrix, where N is number of stocks and T is number of dates, and $N \ll T$. Assume the eigenvalues of AA^{T} is 5, 4, 3, 2, 1, 0,associated with eigenvectors v_1,..., v_N, what are the eigenvalues and eigenvectors of $A^{\mathrm{T}}A$?

2. (35 points) Consider the probability density function

$$f_X(x|\gamma) = \begin{cases} c_\gamma(1 - \frac{x}{\gamma}) & \text{if } 0 \le x \le \gamma \\ 0 & \text{otherwise} \end{cases}$$

where c_γ is a constant that may depend on γ. You can assume that $\gamma > 0$. Let X_1,..., X_n be i.i.d. random variables each with p.d.f $f_X(x|\gamma)$.

   (a) (5 points) Provide a simplified expression for c_γ.

   (b) (10 points) What is the method of moments estimator, $\hat{\gamma}_{MM}$, for γ based on the second moment of X, i.e. X_1^2,..., X_n^2? Is $\hat{\gamma}_{MM}$ an unbiased estimator of γ?

   (c) (10 points) What is the limiting distribution of $\sqrt{n}(\hat{\gamma}_{MM} - \gamma)$?

(d) (10 points) Let

$$\pi(\gamma) = \begin{cases} \frac{\gamma^3}{20} & \text{if } 1 \leq \gamma \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

be a prior distribution for γ. Based on a single observation X_1, what is the mean of the posterior? You may assume $x_1 > 0$.

3. (25 points) Twitter is a social media website where users can follow each other. Suppose a group of N hedge fund managers are all registered on Twitter, and any manager has independent probability p of following another.

(a) (10 points) Alice, Bob and Joel are all enrolled on the website. What is the probability that Alice has the same number of followers as Bob, and both have more followers than Joel?

(b) (15 points) On Twitter, a *manager circle* of length n occurs when we can find managers s_1, ..., s_n, so that s_1 follows s_2, s_2 follows s_3, ..., and s_n follows s_1. Let X_n be the number of *manager circle* of length n. Find the expected value $\mathbb{E}(X_3)$ and variance $var(X_3)$.

(c) (bonus, 15 points) Let's define there is *one connection* between users A and B if either A follows B, or B follows A, or they follow each other. There is one situation that, for any $N-2$ users among all the N managers, there are 3^k connections among those $N-2$ people, where k is a positive integer. Derive all the possible values for N. (Hint: it is a realized situation, so the probability p is not relevant to this question)

4. (25 points) a, b are two sequences of integers of the same length N, $a[i], b[j]$ denotes the i-th number of a and the j-th number of b. Design an algorithm to find the largest subset from $F \subset \{1, 2, \cdots, N\}$ such that for any $k, t \in F$, we have if $a[k] > a[t]$, then $b[k] > b[t]$; or if $a[k] \leq a[t]$, then $b[k] \leq b[t]$.

   If we allow n pairs of (t, k) to break the order requirement, then how to design the algorithm?

ORSUS Research

Programming Exam

Orsus Research Exam

Time Limit: 120 Minutes

Date: ________________

Start Time: ____________

End Time: ____________

Name (Print): ______________________

University: ________________________

Year: ____________________________

Question 1 – [General Programming] (3 points/question, 27 points)

1. What is pass by value vs. pass by reference?

2. What is a heap and a stack in a process? What are they used for?

3. What is the relationship between threads and processes?

4. What is a class? What is an object?

ORSUS Research

5. What kind of environment variables are essential for programming environments?

6. What is the difference between iteration and recursion? When and why would you use one over the other?

7. Compare and contrast a linked list and an array. Provide some use cases for each.

8. How does the compiler find the built-in libraries and packages when import or include keywords are used?

9. What are sets and maps? How do they work internally? Provide some use cases for each.

Question 2 – [SQL] (3 points/question, 18 points)

1. What is a clustered index? What is a non-clustered index?

2. What is a trigger?

3. Compare and contrast the various joins.

4. What is an aggregate function?

5. What is the difference between having and where clauses?

6. What is a stored procedure? What are the benefits of it?

Question 3 – [Implementation] (55 points)

Implement the following functions. Provide space and runtime metrics for all options. Please write/type these functions on separate pages.

1. (5 points) Post order iteration of a binary tree.
   - printPostOrder(Node root)

2. (15 points) LRU (least recently used) cache.
   - getItem(DataObject obj)

3. (5 points) Find nth Fibonacci number
   - getFibonacciNumber(n)

4. (5 points) Depth First Search of Binary tree.
   - findElement(Node root, int element)

5. (25 points) Given a cost matrix (each cell has cost to reach itself), find the minimum cost to reach cell (m,n) from cell (a,b).

   In this matrix, the minimum cost path to reach cell 3,2 is as shown:

3	2	8
1	9	7
0	5	2
6	4	3

   Hence, minimum cost is = 11

ORSUS Research

Finance Exam (Optional)

Orsus Research Exam

Time Limit: 60 Minutes

Date: ________________

Start Time: ____________

End Time: ____________

Name (Print): ______________________

University: ______________________

Year: ______________________

Question 1 – [Regression Analysis] (25 points)

(a) (15 points) What is the relation between the estimated coefficient in a linear regression when: (a) Y is regressed on X, and (b) X is regressed on Y?

(b) (10 points) Highlight the difference and explain the result intuitively, in words. (Hint: Feel free to draw a graph if it is helpful.)

ORSUS Research

Question 2 – [Portfolio Optimization] (75 points)

Consider a universe of N risky assets. The expected return and volatility for asset *i* are denoted by μ_i and σ_i. The correlation coefficient between assets *i* and *j* is ρ_{ij}. The vector of expected returns as the (column) vector $\mu = (\mu_1, \mu_2, \ldots, \mu_N)'$. The variance-covariance matrix is denoted by Σ, where the diagonal components are $\Sigma_{ii} = \sigma_i^2$ and the off-diagonals are $\Sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$.

We want to find the optimal allocation of weights across assets, i.e. $w = (w_1, w_2, \ldots, w_N)'$, such that the expected return of the portfolio μ_p is maximized subject to a portfolio risk target $\sigma_{p,0}{}^2$.

(a) (15 points) Write down the optimal portfolio problem in matrix form.

(b) (15 points) Write down the dual problem in matrix form. (Hint: Feel free to use $\mu_{p,0}$ as the expected return on the global minimum variance portfolio.)

(c) (20 points) Express the Lagrangian and the resulting optimal conditions for the dual. How many equations do you have? What are the unknown variables?

ORSUS
Research

(d) (15 points) Consider the case of 2 assets, i.e. *N=2*. Let $\mu_1 = 0.02$, $\sigma_1 = 0.02$, $\mu_2 = 0.08$, $\sigma_2 = 0.15$, and $\rho_{12} = 0$. What are the optimal weights for a portfolio with expected return of 0.05? What is the overall risk and Sharpe ratio (SR= μ_p/σ_p) of this portfolio?

E) (10 points) What happens when the correlation between assets increases to $\rho_{12} = 0.5$. Explain what happens intuitively.