A special issue on Data Science

FEATURES
Deblurring the unknown
Clearer brain scan images
Mathematics in the deluge of DNA data
Making every child death count
The Faculty of Science conducts basic and applied experimental, theoretical and simulation research over a broad spectrum of science, mathematics and technology domains. We cover most of the key fields in biological sciences, chemistry, physics, pharmacy, mathematics and statistics.

Faculty of Science Research is published twice a year. It is written for a broad scientific audience interested to keep up with some of the key areas of science pioneered by researchers at the Faculty of Science.

This publication may be reproduced in its original form for personal use only. Modification or commercial use without prior permission from the copyright holder is prohibited.

Correction
In the Jun 2014 “A lifetime of contributions to Science”, it was wrongly reported that Professor Hardy CHAN had been with NUS since 2000. Professor CHAN had in fact been with NUS since 1981. We apologise for the error.

On the cover: Data science comprises science-driven data research and data-driven science research. Together, these provide the innovations required to collect, collate, manage, process, transform and store massive data sets, called Big Data.

FOS RESEARCH | VOL. 19 | NUMBER 2 | DECEMBER 2014
News Roundup

Charles Darwin’s Beagle library reconstructed

The voyage of the Beagle around the globe between 1831-1836 is one of the most important scientific expeditions in history. On board was the young Charles Darwin. His investigations on the Galapagos and elsewhere would change science and the world forever. The Beagle may not have had the internet, but she did have a stunning state-of-the-art library during that time. The library was dispersed at the end of the voyage. Its contents long remained a mystery. Now, 178 years later, the library has been reconstructed as part of Darwin Online (http://darwin-online.org.uk) by historian of science Dr John VAN WYHE from the Department of Biological Sciences.

The library consists of 404 volumes amounting to over 195,000 pages and was reconstructed by meticulous research for clues in surviving documents from the voyage. Perhaps the most impressive revelation in the online Beagle library is its rich visual gallery of thousands of images.

The library played a major part in Darwin’s research, serving as reference and inspiration for the voracious reader and inquisitive naturalist. Having the library available online allows us to obtain new insights into his landmark expedition around the world that ultimately resulted in the creation of modern evolutionary biology.

Dr VAN WYHE said, “Darwin lived and worked in the Beagle library for five years. The online library provides an unprecedented insight into the journey that changed science and our understanding of the world.”

Therapeutics to fight cancer

The biotechnology research group led by Professor WANG Shu from the Department of Biological Sciences is collaborating with FF CanVac, a Singapore-based biotech company to develop new anti-cancer therapeutics which can train a patient’s T cells (T lymphocytes) to attack cancerous tumours. Using the patient’s own immune cells to recognise and destroy cancer cells, so called cancer immunotherapy, is a technique that was hailed as Science magazine’s breakthrough of 2013. FF CanVac is a clinical-stage company specialising in the research, clinical development, and commercialisation of cancer immunotherapy treatments. This collaborative project seeks to translate the scientific discoveries from Professor WANG’s lab to the clinical trials stage and facilitate the development of commercially viable treatments to improve health outcomes.

Immunotherapy is set to revolutionise the treatment of cancer, and has advanced with great strides in the past decade. Various forms of cancer immunotherapy products have recently been demonstrated to produce dramatic responses in cancer patients who had failed conventional modalities like chemotherapy and radiation therapy. Professor WANG commented, “Cancer immunotherapy will one day become a significant form of treatment to fight against various late-stage cancers.”
Green and sustainable manufacturing

The research group led by Professor HUYNH Han Vinh from the Department of Chemistry, together with a team led by Dr DUONG Anh Hung from the Institute of Chemical and Engineering Sciences, Agency for Science, Technology and Research (A*STAR) are collaborating with global pharmaceutical company, GlaxoSmithKline (GSK) to develop new synthetic technologies to replace transition metal catalysts with novel iron (Fe) based ones. This research is supported by the “GSK-Singapore partnership for green and sustainable manufacturing” initiative.

Numerous pharmaceuticals, agro-chemicals and materials feature biaryl scaffolds as important building blocks in their structures, which in their simplest forms consist of two aromatic rings directly linked to each other. Transition metal based catalysts are used in the production of these building blocks by coupling two aromatic rings with each other. At present, biaryls are prepared using palladium or nickel containing catalysts. Many of these heavy metal compounds are costly and often toxic. Iron, which is abundant in nature, is a possible low cost and environmentally more benign alternative. The successful realisation of this research could improve the sustainability of pharmaceutical and fine chemical manufacturing and potentially mitigate some major industrial challenges, such as toxic waste removal. Professor HUYNH remarked, “Working toward sustainability should be a common goal for everyone. Organometallic chemistry, in particular, plays a major role as it can provide valuable tools in this endeavour.”

Improving survivability in kidney transplant patients

The research team led by Dr YAU Wai Ping from the Department of Pharmacy, together with the group led by Professor Anantharaman VATHSALA from the Department of Medicine, are supported by the National Kidney Foundation (NKF) in a joint effort to study the impact of genetic variants on the safety profile of immunosuppressive agents in Asian kidney transplant patients.

Calcineurin inhibitors (CNIs), which include medications such as cyclosporine and tacrolimus, are usually administered to kidney transplant patients to prevent rejection by the body. Although CNIs improve short-term transplant survival by reducing acute rejection rates, they are associated with adverse effects including new-onset diabetes after transplantation (NODAT). NODAT is a major metabolic complication that increases risks of kidney graft failure and reduces patient survival. Differences in the incidence of CNI-associated NODAT among different racial and ethnic groups have been observed clinically and genetic factors could be a possible reason. Findings from this research will provide insights into the genetic predisposition for CNI-related NODAT in Asian kidney transplant recipients, and aid in the identification of potential genomic biomarkers for assessing the risk of developing NODAT.

Putting into perspective the clinical significance of this work, Dr YAU explained, “The knowledge gained from this study may aid clinicians identify high-risk patients to whom immunosuppressive regimens could be tailored to prevent NODAT development and improve long-term transplant survival in Asian kidney transplant population.”
Deblurring the unknown

Innovative mathematical techniques can restore missing information in photographs, leading to clearer images

Introduction

One challenging problem often seen in imaging sciences is the so-called blind deconvolution problem, which aims at recovering the clear image from one blurred observation without knowing the blurring process. Built upon the concept of sparse modelling/approximation, several new mathematical frameworks and computational techniques that are developed provide a strong theoretical foundation, as well as effective computational methods, for resolving the blind deconvolution problem. Moreover, these techniques can also be used for solving many other inverse problems in imaging sciences.

Despite the astonishing progress in digital imaging technology, the demand of higher image quality at lower hardware cost is still growing rapidly in many imaging applications. Image recovery aims at restoring images which are of better quality from input images in which important information might either be severely attenuated or are completely missing. Image recovery usually is an ill-posed inverse problem such that a direct reversing process cannot recover missing information and will amplify image noise such that the recovered image is severely distorted. Even worse, in many imaging applications, how input images are blurred is not known or cannot be accurately estimated by hardware. Thus, the problem of recovering a clear image from a blurred image is called a blind deconvolution problem, which is often seen in digital photography and scientific imaging.

Image Blurring

Image blurring is one primary cause of low-quality images. Image blurring can be modelled as a convolution process which will eliminate or severely attenuate high-frequency component of images. Image de-blurring is about recovering the missing information and restoring attenuated information. De-blurring is an ill-posed inverse problem as a direct reversing process cannot recover the missing information and will amplify image noise such that the recovered image is severely distorted. Even worse, in many imaging applications, how input images are blurred is not known or cannot be accurately estimated by hardware. Thus, the problem of recovering a clear image from a blurred image is called a blind deconvolution problem, which is often seen in digital photography and scientific imaging.

Blind Image Deconvolution

Blind image deconvolution is a challenging problem with many solutions that are sound in Mathematics. One obnoxious solution is having a blurred image which is generated by itself and a trivial convolution process that does nothing. A fundamental question is how to define sharp images and how to estimate them from blurred ones. The author and his collaborators have developed several new mathematical tools and computational frameworks for resolving this challenging problem. The basic idea is to model sharp images and the kernels that determine the blurring process by sparse approximation, which assumes that the data of interest are compressible in some suitable transform domain. For a sparse approximation based blind image deconvolution system, there are three main components:

1) what transform can sparsely model data for estimation;
2) how to find the sparse representation of data under the given transform; and
3) how to solve the problem in a robust manner that takes all kinds of practical error sources into account.

Wavelet Tight Frames

Built upon the concepts of wavelet tight frames and sparse approximation via $\ell_1$-norm relating optimisation, the author and his collaborators developed in [1] a powerful sparse approximation based framework for blind image deblurring. The proposed method alternatingly estimates the blur kernel and the sharp image via solving an $\ell_1$-norm relating optimisation problem. Wavelet tight frame is used for sparsely approximating the image, as well as the blur kernel. In a subsequent work, it is observed that the $\ell_1$-norm based sparse approximation is biased toward a slightly blurred result, and thus the ratio $\ell_1/\ell_2$ is proposed as the normalised sparsity prompting function for correcting the bias of the result using $\ell_1$-norm based sparsity-prompting function. The proposed computational framework can effectively restore a large class of blurred images taken by hand-held cameras. Some of these techniques have drawn attention from the industry and some companies have assigned consultation contracts for the implementation of these published algorithms in their camera systems to improve image quality.
JI Hui received his Ph.D. in Computer Science from University of Maryland at College Park in 2006 and joined the Department of Mathematics at NUS as an assistant professor. His primary research interest is applied and computational harmonic analysis with a particular focus on developing mathematical theories and methods to solve practical problems in the field of computer vision and imaging sciences. He is also director of the Centre for Wavelets, Approximation and Information Processing (CWAIP).

References


**Image Deconvolution**

The convolution process for modelling image blurring is often a simplified model which does not exactly reflect the physical process happening in practice. A high-performance non-blind or blind deconvolution system that works well in practice needs to take such a model error into consideration. Built upon a new sparse approximation based variational model, a robust non-blind image de-convolution technique is developed in [2], the first available technique that is robust to model error. The basic idea is to simultaneously estimate three components:

1) sharp image;

2) the distortion of image gradients; and

3) the artifacts caused by kernel error.

All three components can be sparsely approximated under different transforms. The proposed method is applied on the non-stationary blind image deconvolution problem and it has showed superior performance over the existing techniques.

**Data-driven Wavelet Tight Frames**

One key for further improving the performance of state-of-the-art sparse approximation based image recovery method is to design new transforms that can be more efficient on sparsely approximating input images. Built upon the concept of “adaptive learning”, a new approach is developed in [3] which constructs a class of data-driven wavelet tight frames that are optimised for the input data, instead of the existing generic wavelet tight frames. The data-driven wavelet tight frames constructed from the proposed approach can effectively capture the structures of input image data, which leads to better performance of sparse approximation to input image. Such an improvement over existing generic wavelet tight frames leads to noticeable performance gain in many wavelet based image recovery methods.

The techniques developed for blind image deconvolution and data-driven wavelet tight frame not only can be used for solving the related image recovery problems, but also can be applied to solve other inverse problems arising from imaging sciences, including 3D reconstruction in electronic microscopy and computed tomography, signal processing for high-dimensional data and many others.
Clearer brain scan images

A novel method to improve the visual clarity of brain images in an efficient way

Introduction

Diffusion tensor imaging (DTI) has been widely adopted by neuroscientists to investigate the anatomical structure of human brains, since it can characterise the diffusivity behaviour of water at each volume element (voxel) over a 3D imaging space and consequently connects the microscopic properties of the brain tissue to the macrostructure of the white-matter fibers. For each brain voxel, DTI characterises the water diffusion property through a diffusion tensor (DT), which is mathematically represented as a 3×3 symmetric positive definite matrix. On the other hand, DTs are estimated from the diffusion weighted imaging (DWI) data acquired from magnetic resonance experiments. The raw DWI data, however, carry significant noise; this contaminates the DT estimates, introduces systematic bias into the induced eigenvalues and consequently affects the downstream DTI based scientific studies such as fiber area identification and fiber tracking. Developing efficient methods for accurate DT estimation becomes more important, imminent but challenging; see for example [1, 2] and the references therein.

Recently, a two-stage spatial shrinkage estimation (SpSkE) procedure was proposed by [3] to estimate these DTs efficiently. This method appropriately accommodates the spatial information carried in DWI data and reduces the bias components in the corresponding derived eigenvalue estimates. This work developed the necessary theory and applicable algorithms that are ready for practical usage.

Associations Among Fiber Structures, DT Morphologies and Eigenvalue Relations

A set of DTI data is typically composed of DTs, each contained in a brain voxel. Each DT is mathematically represented as a 3×3 symmetric positive definite matrix, and can be further decomposed to obtain its eigenvalue-eigenvector pairs, \((\lambda_i, \mathbf{v}_i)\), where \(\lambda_i \geq \lambda_j \geq \lambda_k\) for \(i = 1, 2, 3\), which respectively characterises the degree and orientation of water diffusion in the corresponding voxel. Such decomposition can be intuitively illustrated as a 3D ellipsoid; see Figure 1.

Figure 1: Ellipsoid representation of a DT.

With aforementioned decomposition, \(\lambda_{(1)}\) and \(\lambda_{(2)}\) determine the DT morphology, and subsequently with the eigenvectors imply the fiber
structures. In particular, when $\lambda_{12}, \lambda_{22}=0$ and $\lambda_{11}=\lambda_{12}=0$, the DT is approximately spherically shaped, meaning that there is no fiber passing through this voxel; in this case the corresponding eigenvectors are arbitrarily oriented. When $\lambda_{11}, \lambda_{22}>0$, the DT is either prolate ($\lambda_{11}>0$, the fiber is single oriented or crossed with a small angle) or nondegenerate ($\lambda_{12}, \lambda_{11}>0$, two or more fibers are crossed with different intensities on different orientation). The eigenvector associated with $\lambda_{11}$ is then pointing at the orientation of the single oriented fiber tracts or the fiber tracts with greater intensities. When $\lambda_{11}, \lambda_{12}=0$ but $\lambda_{12}, \lambda_{11}>0$, the DT is obtuse. This usually corresponds to crossed fiber tracts with similar intensities on two more different orientations; the first two eigenvectors are arbitrarily located in the plane perpendicular to the eigenvector corresponding to $\lambda_{11}$. An illustrative display of the associations among the fiber structures, DT morphologies, and eigenvalue relations is given in Figure 2.

Spatial Shrinkage Diffusion Tensor Estimation

Classical methods in the literature estimate the DT in every voxel independently. The DT estimates based on these classical approaches are severely contaminated by the high noise of the DWI data commonly encountered in practice and introduce significant systematic bias into the corresponding eigenvalue estimates. Therefore, statisticians in the community proposed a number of approaches to improve the DT estimation. Examples include robust estimation procedures, Gaussian estimation procedures, Gaussian weighted least-square function. The locally linear model, SpSkE incorporates the framework of the heteroscedastic linear model (HLM) method, and structural adaptive smoothing. However, as discussed in [3], all these methods do not simultaneously hold the following properties:

1. appropriately accounting for the spatial information carried in DWI data, since the DWI data are typically spatially structured;

2. correctly addressing the bias problem in the eigenvalue estimation;

3. robust with respect to the noise coming from different magnetic resonance experiments; and

4. establishing sufficient theoretical foundation for the proposed method to support its practical application.

In [3], a two-stage SpSkE method was proposed; it has all the good properties mentioned above. In particular, under the framework of the heteroscedastic linear model, SpSkE incorporates $L_1$-type penalisation and the locally weighted least-square function. The main idea is:

1. using the neighbourhood information to refine the DT estimation; and

2. incorporating the global information to help reduce the bias component in the eigenvalue estimation.

With a two-stage structure, the algorithm first refine the DT estimation (stage 1) and then look into the decomposed DT structure and focus on the eigenvalue estimation and its bias reduction (stage 2).

Data Application

The DT estimates for all brain voxels are obtained by applying the SpSkE method to a real DWI data example. The results are then compared with two existing methods in the literature: the WLS method by [1] and the StAdS method by [2]. The colour-coded fractional anisotropy (FA) maps for one brain slice from these three methods, as given in Figure 3. It can be seen from this figure that SpSkE very effectively reduces the FA values in those areas highly likely to be non-fiber areas of the brain. Therefore, it is expected to effectively alleviate the confounding information carried in the DT estimates and assist the subsequent fiber-tracking processes by reducing the faulty findings.

YU Tao received his B.S. degree and M.S. (supervised by Professor Shiyi SHEN) in Mathematics and Probability & Statistics from Nankai University in 2001 and 2004 respectively. He obtained his Ph.D. under the direction of Professor Chunming ZHANG from University of Wisconsin-Madison in 2009. After that, he was appointed assistant professor with Department of Statistics and Applied Probability, NUS.

References


Mathematics in the deluge of DNA data

A more efficient way of decoding DNA information using spaced seeds

Introduction

Human beings are evolutionary products of nature and environment. Here, the nature is the DNA and the environment includes food, air and many other inputs to individuals. To fully understand ourselves, we need to know our DNA, a string of three billion nucleotides of four kinds, abbreviated as A, G, C, and T. Human cells contain two copies of the DNA, one inherited from the father and one from the mother. The whole DNA (that is genome) is broken up into 23 “modules”, called chromosomes, and within each chromosome are sparsely scattered genes. Each cell uses the genes to determine which proteins should be made to perform specific cellular functions (such as metabolic reactions and transducing chemical signals) at a particular time point. The Human Genome Project (http://www.genome.gov) spent $2.7 billion on completing the first draft of the human genome from 1991 to early 2000.

Challenges In Decoding DNA Information

Decoding DNA information creates formidable computational challenges in data storage, transfer and analysis. Thirty years ago, there were little DNA sequences and computers were slow. Presently, computational speed is improving but at a rate much less than sequencing machines. Under Moore’s Law, computer processors doubled in speed every 18 months. In comparison, between 2008 and 2014, DNA sequence data has grown about three- to fivefold per year.

Change (mutation) in DNA never stops in living organisms. A mutation may delete a nucleotide, insert a nucleotide, or replace one nucleotide by another. Large-scale mutations also occur, which lead to one segment of DNA duplicated or relocated at a different position, from time to time. Mutation leads to gain and loss of genes and can affect the fate of a living organism during evolution.

Since DNA has a hidden structure, it is traditionally analysed in a comparative manner. For example, when the genomes of different species are compared, one will identify the conserved regions where different species have almost identical nucleotide sequences and the volatile regions where each species has a different nucleotide composition. Conserved regions are likely genes and other interesting functional units that a species cannot afford to lose. Therefore, the sequences of a conserved region in different species which descend from a common ancestor are called homologs. Since convergent evolution that leads different genes to have the same sequence occurs rarely, highly conserved regions are sought to infer genes or other interesting functional units.

Mathematics Behind Homolog Search Tools

When genomes are compared, most regions are dissimilar. Any exact algorithm for homolog search is time-consuming, as it wastes a huge amount of time in dissimilar regions. Consider human and mouse genomes, each having over three billion nucleotides. An exact algorithm may take years to compare them!

Presently, widely-used computer programs for homolog search are all designed based on the so-called filtration approach. To determine potential locations of conserved regions in the query DNA sequences, a filtration-based program first records down the occurrences of all 4i words of a fixed length l (say 11) in each sequence. A pair of positions (k, j) is called a seed hit if there is a word of length l that occurs in the k-th position in one sequence and the j-th position in another. The program then extends each seed hit in both directions to look for long conserved regions. Since the number of seed hits reported in the first step is significantly small, compared to all possible pairs of positions, a filtration-based program improves in speed by an order. Such a program usually takes 10 to 20 days to compare human and mouse genomes on a desktop computer.

What about the performance of the filtration-based programs? Clearly, the filtration approach sacrifices sensitivity for speed. The larger the word length l is, the less seed hits are reported in the first place and hence the more homologous sequences are likely missing.

Seed Selection

Next, how to select a seed? The BLAST program (http://www.ncbi.nlm.nih.gov/) is the official homolog search program for searching DNA and protein sequence databases at the National Center for Biotechnology Information, USA. When it was launched in 1990, the BLAST used a perfect word match of length 11 as a seed hit, meaning 11 consecutive identical nucleotides found in the two input sequences will trigger the extension process. Requesting 11 consecutive base matches seems a good choice. However, mathematics analysis suggests otherwise! That is, it is much better to use identical nucleotides in a set of non-consecutive positions, called a spaced seed, to trigger the extension process. For example, we may request to have identical nucleotides in the 1\textsuperscript{st}, 2\textsuperscript{nd}, 4\textsuperscript{th} and 6\textsuperscript{th} positions in a window of length 6. Such a spaced seed is written...
ZHANG Louxin received his Ph.D. from the University of Waterloo and is currently an associate professor at the Department of Mathematics, NUS. His research mainly focuses on building mathematical models and designing algorithms that allow biologists to better analyse DNA sequences and to decipher gene regulation.

References
Making every child death count

Evaluating global progress in reducing child mortality using new statistical methods

Introduction

Substantial progress has been made in reducing child mortality globally since 1990. The under-five mortality rate, which is the probability that a child born in a given year will die before reaching the age of five, has decreased globally from 90 deaths per 1,000 births in 1990 to 46 deaths per 1,000 births in 2013 [1]. While this translates into a decline in the number of under-five deaths from 12.7 million in 1990 to 6.3 million in 2013, it still implies that about 17,000 children died before their fifth birthday on each day in 2013. Monitoring of levels and trends in child mortality is necessary to further improve child survival and to end preventable child deaths.

National estimates of child mortality are produced by the United Nations Inter-agency Group for Child Mortality Estimation (UN IGME) and used to track progress in reducing child mortality. To construct child mortality estimates, the UN IGME compiles data from various sources, typically from vital registration systems, surveys and censuses. The data sources either record recent births and deaths on an ongoing basis or collect retrospective information on child mortality in the form of full or summary birth histories of women. Estimating levels and trends in child mortality is challenging for many developing countries without well-functioning vital registration systems because of limited availability of recent data and the occurrence of data errors, for example random errors in sample surveys or systematic errors due to misreporting, as illustrated in Figure 1 for Nigeria. To overcome these challenges and produce estimates for all populations of interest, the UN IGME uses a variety of demographic and statistical methods. This article gives three examples of such methods, which were recently developed by Professor ALKEMA and her collaborators to provide new insights into child mortality at the national and subnational level. Each of the methods is currently in use by the UN IGME and more detailed findings are given in [1].

Estimating Child Mortality For All Countries

To assess levels and trends in the under-five mortality rate, Professor ALKEMA and NUS graduate Jin Rou NEW developed the Bayesian B-spline Bias-reduction model [2]. In this model, a smooth curve is fitted to the available data in a country. In the fitting process, information on the smoothness of USMR trajectories is exchanged across countries to be able to obtain estimates and projections for countries with limited data. To account for data quality issues, an observed value for USMR is considered as the true value for USMR multiplied by an error factor. Differences in the expected values of the errors (or biases) as well as the differences in uncertainty in errors (or variances) are accounted for when constructing USMR estimates through the inclusion of a data quality model. The resulting model produces estimates for all countries with uncertainty. Illustrative estimates for Nigeria are shown in Figure 1.

Assessing Differences In Child Mortality Between Boys And Girls

Under natural circumstances, girls tend to have advantages over boys with respect to survival up to age five, resulting in a sex ratio of male to female USMR which is greater than one. This survival advantage for girls tends to increase as total USMR (for both sexes combined) decrease, explained by changes in the associated cause of death distributions, which are generally more favourable for girls’ survival at lower mortality levels. However, gender discrimination can result in unusually high or low sex ratios, for example through disadvantaging treatment of girls as compared to boys, as documented in various, mostly Asian, countries.

To estimate gender-specific mortality, Professor ALKEMA, NUS graduate Fengqing CHAO and others developed a statistical model in which country-specific sex ratios are modelled by the product of the expected “natural” sex ratio and a country-year-specific multiplier [3]. The expected sex ratio follows from the global relation between sex ratios and total mortality rates, while the multiplier represents the relative advantage or disadvantage of girls to boys compared to other countries at similar total mortality rates. The estimated global relationship is shown in Figure 2a and shows that survival chances tend to improve more rapidly for girls as compared to boys as total mortality decreases, with a reversal of this trend at very low total USMR. For many countries, gender
Leontine ALKEMA is an assistant professor in the Department of Statistics and Applied Probability and the Saw Swee Hock School of Public Health. Her research focuses on developing and validating statistical methods for estimating and projecting population and health indicators in developing countries. She received her Ph.D. in Statistics from the University of Washington and was a post-doctoral fellow at Columbia University before joining NUS in 2009.

References
Pushing research boundaries forward

Research institutes and centres within the University undertake multi-disciplinary research work in areas of excellence. Many of the data scientists at the Faculty of Science have established strong working relationships with them to push boundaries and accelerate scientific discovery. Below are four of these research institutes and centres.

Institute for Mathematical Sciences

The Institute for Mathematical Sciences (IMS), established in 2000, is an internationally recognised institute that fosters fundamental and multi-disciplinary mathematical research to facilitate the growth of mathematical expertise among researchers in Singapore and the international community. To fulfil these aims, IMS organises annually five to six research programmes which incorporate workshops and tutorial sessions. Acting on the recommendations of the institute’s scientific advisory board of eminent scientists, the programmes are chosen from forefront areas of mathematical sciences. In addition, IMS hosts research conferences and summer schools for graduate students; and organises public and school lectures to raise public awareness of the role of mathematics in science, engineering, technology and industry.

For details, please visit http://www2.ims.nus.edu.sg/

Risk Management Institute

The NUS Risk Management Institute (RMI) was established in 2006 as a research institute dedicated to the area of financial risk management. The institute is a globally recognised knowledge centre in financial risk management through research, education and training. Besides the Master of Science in Financial Engineering degree and a series of executive programmes, RMI runs a non-profit Credit Research Initiative (CRI) that “transforms big data into smart data”. Using the probability of default (PD) model as its foundation, it provides various credit risk measures including daily updated PDs and Actuarial Spreads (AS) for individual firms and their aggregations for different sectors, economies and regions, with a current coverage of 116 economies and over 60,000 listed firms.

For details, please visit http://rmi.nus.edu.sg/

Centre for Quantitative Finance

The Centre for Quantitative Finance (CQF) within the Faculty of Science aims to advance knowledge development in quantitative finance. As a leading hub in both research and education in the theory and practice of quantitative finance, the CQF brings together complementary expertise across NUS, particularly from Sciences, Business, and Economics. Researchers at the centre actively engage financial professionals to explore current issues faced by financial markets. The centre also provides short-term courses and programmes to equip working professionals with the latest industry practices.

For details, please visit http://cqf.nus.edu.sg/

Centre for Wavelets, Approximation and Information Processing

The Centre for Wavelets, Approximation and Information Processing (CWAIP) is a research centre of the Faculty of Science. It is a multidisciplinary research centre that emphasizes the synergy of mathematics, engineering and computer science. The main objective of the centre is to push forward the frontier of the areas of visual information processing, analysis, understanding, and apply state-of-the-art mathematical theory and computational technology to solve real-life problems, such as image/video restoration and enhancement, visual recognition, and inverse problems in bio-imaging sciences. The centre also actively supports the educational goals of the University through the supervision of final year undergraduates, postgraduate research students and overseas students on internship.

For details, please visit http://www.cwaip.nus.edu.sg/