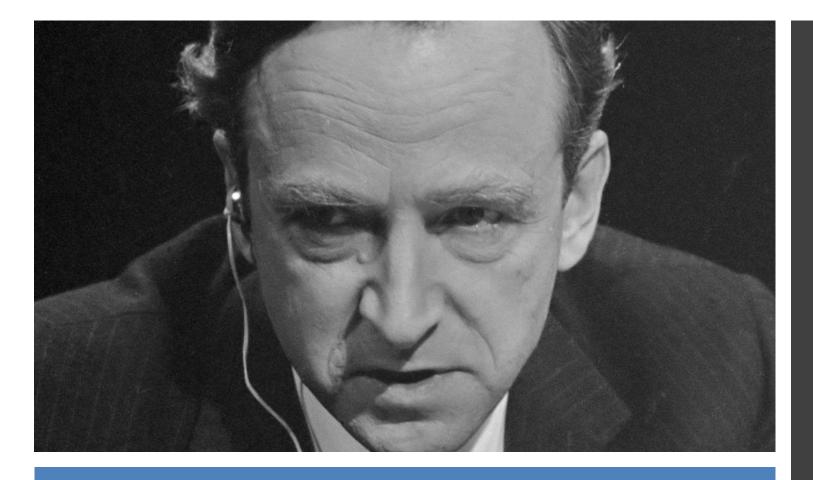
Markov chains Directed Factor Graph Dynamic Bayes nets Graphical Directed Bayesian Models Networks variable Influence diagrams Chain Graphs Strong Undirected Graphs Markov network input dependent CRF Clique Graphs Pairwise machine (disc.) Junction Gauss. Clique **Process** (cont)

Connections, correlations, causality & coherence

Probabilistic graphical models – a framework for sparse prediction and coherent stress testing, in Canadian capital markets

How I stopped worrying & learned to love uncertainty.



In memory
Robert Neild — 1924-2018

- British Keynesian economist who influenced Labour policy and whose other fields of interest included oysters and their aphrodisiac powers
- I remember a damascene moment when, in early 1974 (after playing round with concepts devised in conversation with Nicky Kaldor and Robert Neild), I first apprehended the strategic importance of the accounting identity

– Wynne Godley.



We need to put aside the longestablished Newtonian vision of a harmonious economy with negligible innovations

Abstract

Two key objectives of macroprudential modelling are to measure

interconnectedness between financial products, benchmarks & institutions, from public data, and

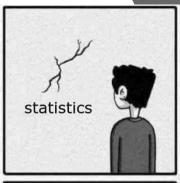
resilience of the system by conducting coherent stress tests.

A tool that has emerged from the machine learning community is (Bayesian) **probabilistic graphical models** which offer algorithms:

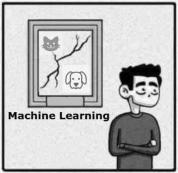
Information filtering networks - parsimonious, performant

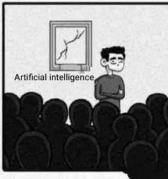
Consistent treatment of probabilities (frequentist vs. subjective) & expert domain knowledge, intuitive

ML for economics









• **Bayesian networks** are to probability calculus what spreadsheets are for arithmetic.

-- Conrady and Jouffe, 2015

 Currently, Bayesian Networks have become one of the most complete, selfsustained and coherent formalisms used for knowledge acquisition, representation and application through computer systems.

-- Bouhamed, 2015

• The core idea at the heart of **model-based machine learning** is that all the assumptions about the problem domain are made explicit in the form of a model. In fact, a model is just made up of this set of assumptions, expressed in a precise mathematical form. These assumptions include the number and types of variables in the problem domain, which variables affect each other, and what the effect of changing one variable is on another variable.

- Winn and Bishop 2018

A good player goes where the puck is. A great player goes where the puck is going to be. - Wayne Gretzky

Bayesian graphical models

parsimonious interconnectednesscoherent stress testing

Connections

- Amplification transmission
- E.g. for products
- Filter strong from spurious –
 Information filtering networks cf GLASSO
- Clustering

Correlation

- Given **historical data**, standard measure of interconnectedness
- Graphical visualization e.g. MST → insight
- Association

Conditional

- Conditional probabilities on graph define ML model
- Conditional expectations regression, prediction ("supervised learning")

Coherence

- Coherent stress tests
- Combine frequentist & subjective probs
- And expert knowledge esp. cause
 & effect

Causality

- Correlation does not imply...
- See. Do. Imagine.
- Cognitive
- Why?

Systemic risk & models

Definition of systemic risk

Phases of a crisis

Systemic Risk

Systemic risk related to capital markets

3. In this Act, systemic risk related to capital markets means a threat to the stability of Canada's financial system that originates in, is transmitted through or impairs capital markets and that has the potential to have a material adverse effect on the Canadian economy.

MR: Market risk

CR: Credit risk

FSI: Financial Soundness Indicators

T-model: Threshold Model

• DSA: Debt Sustainability Analysis

• CCA: Contingent Claims Analysis

BSA: Balance Sheet Approach

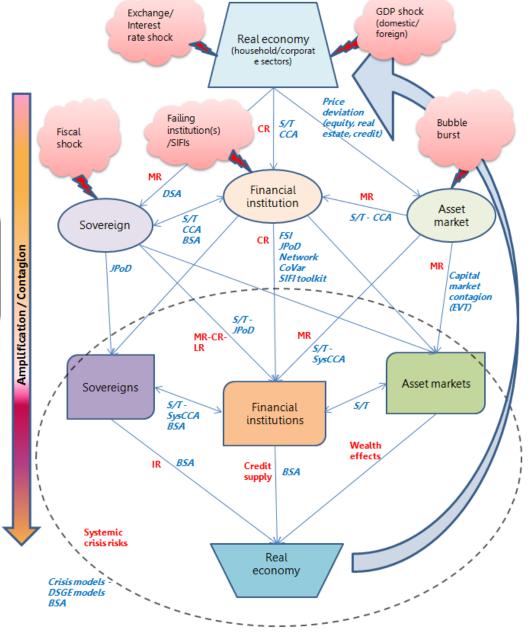
JPoD: Joint probability of default

S/T: Stress testing

EVT: Extreme value theory

http://ccmr-ocrmc.ca/wp-content/uploads/cmsa-consultation-draft-revised-en.pdf

Source: SysMo toolkit (IMF)



<u>Legend</u>: MR: market risk; IR: Interest rate risk; CR: credit risk; DSA: debt sustainability analysis; S/T: stress testing; CCA: Contingent Claims Analysis; SysCCA: Systemic CCA; FSI: Financial Soundnes indicators; JPoD: Joint probability of Default; EVT: Extreme Value Theory

Macropru oversight through a crisis

Macroprudential oversight	Phases of crisis	Model / Indicators	Examples
Risk identification Vulnerabilities Triggers	I. Build-up of vulnerabilities & imbalances	Early-warning (EWI)Financial soundness (FSI)Market intelligenceBacktesting	 Indebtedness of governments, sectors, institutions, households
 Risk assessment Transmission channels Severity of losses \$ Resilience Likelihood	II. Exogenous aggregate shocks	 Macro stress-test (ST) on resilience of financial system Risks: market, credit, liquidity 	MFRAF ST
	III. Amplification and propagation	 Contagion & spillover (FIs, sectors, countries) Impact to real economy 	 Battiston et al: DebtRank Duarte & Eisenbach: Aggregate Vulnerability fire-sale spillover Cont
Risk mitigation	Address any or all of the 3 phases	Impact assessment of proposed measures	See above

Source: Sarlin 2014

Cjupyter Jupyter notebooks

Let us change our traditional attitude to the construction of programs: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

Donald Knuth



☐ The Jupyter Notebook is an open- source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text.
□Render
☐ Github
☐ Nbviewer - https://nbviewer.jupyter.org

□Host

☐ Binder

■ Anaconda

☐ Azure Notebooks

https://notebooks.azure.com/

☐Also: JupypterLab





Model demos







https://notebooks.azure.com/ian-buckley/projects/systemic-risk

- Filtering **R**
- CrisisMetrics*
- Azure ML Studio*

- CoVaR
- SRISK RPy

Resiliency



- Entropy NRM **■R**
- Bayesian■ R
- CIMDO Py
- Corr MST FNA*

Network construction





- DebtRank NRM
 - **■R**
- NEVA Py
- igraph

Network analytics

- ABM-OFR NetLogo
- ABM-ABCE Py
- SFC Py
- Keen Minsky Py

Economic



- PGMs & PP
 - PyMC3
 - Infer.NET
 - DoWhy

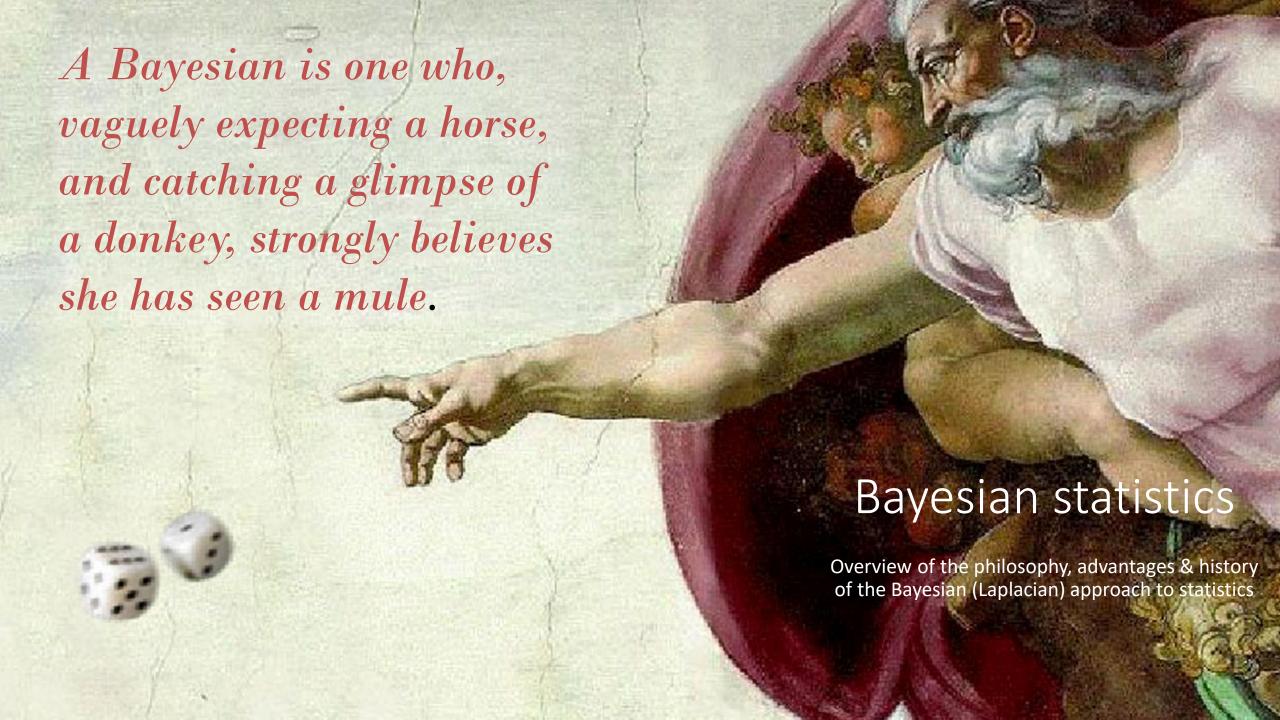
AI/ML



- Corr MST FNA*
- KeyLines*
- Ndtv ■R

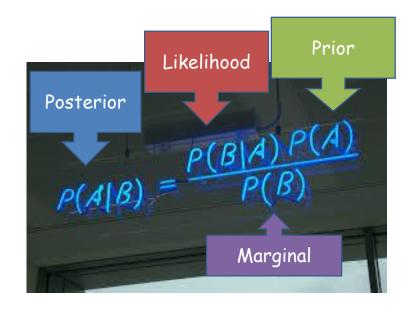
Visualization

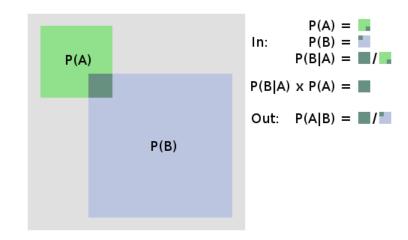




Bayesian vs Frequentist

	Frequentist	Bayesian
Inputs	Data	Data + prior
Uncertainty measurement	Confidence interval	Credible interval
Assess significance	Null hypothesis tests $p(y H_0)$	Direct interpretation of the posterior $p(\theta y)$
Basic concept	Relative freq of an event	Bayes theorem
Probabilities	Inherent property of random phenomena in long run	Plausible knowledge
Parameter estimate	Point	Distribution / sample



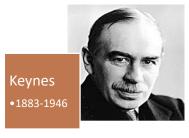


Innovators

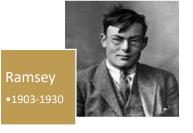
Bayesians





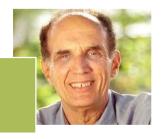










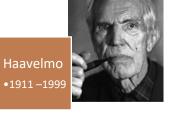




Causality







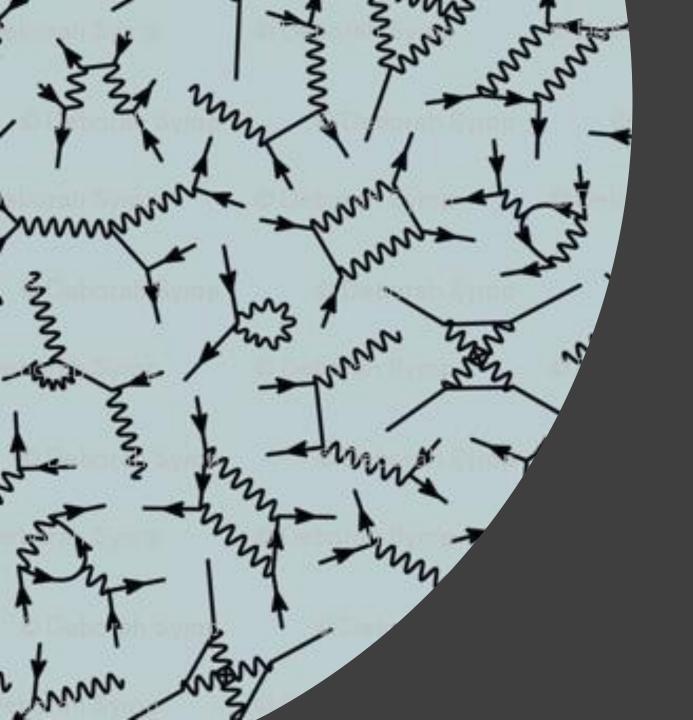








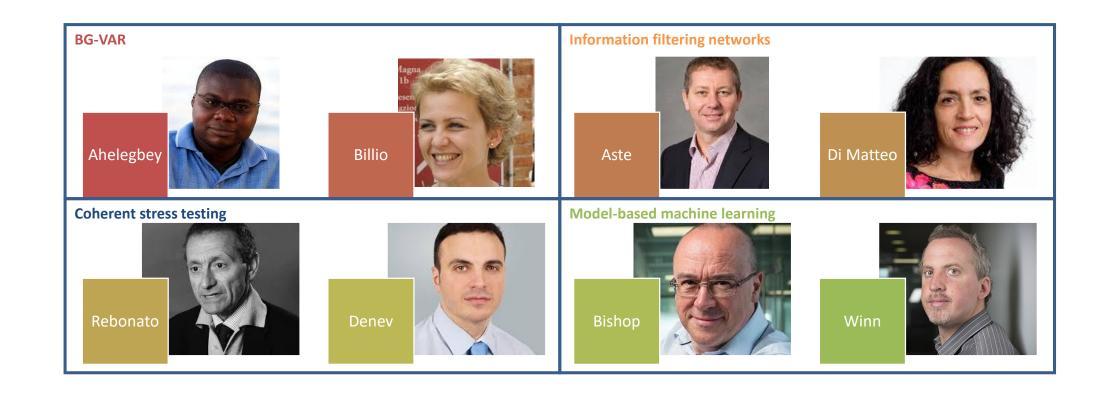




Probabilistic graphical models

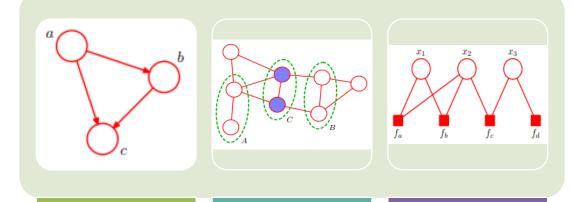
Diagrammatic representations of probability distributions

People – Bayesian graphical models



Graphical models

- A graph comprises
 - **nodes** (*vertices*) connected by
 - links (edges or arcs).
- In a probabilistic GM, each
 - node represents a random variable
 - links express probabilistic relationships between them



Bayesian networks

- Directed DAG
- causal

Markov random fields

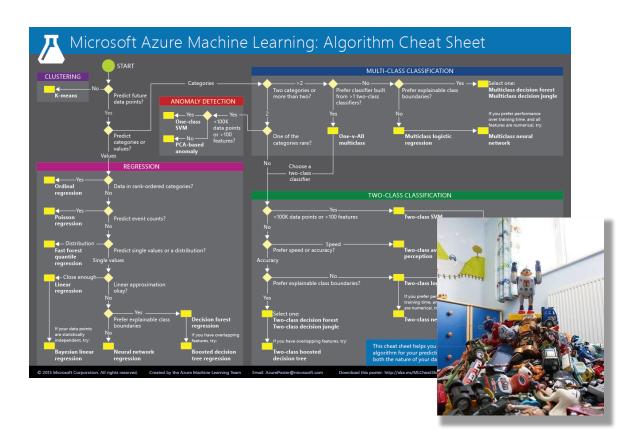
undirected

Factor graphs

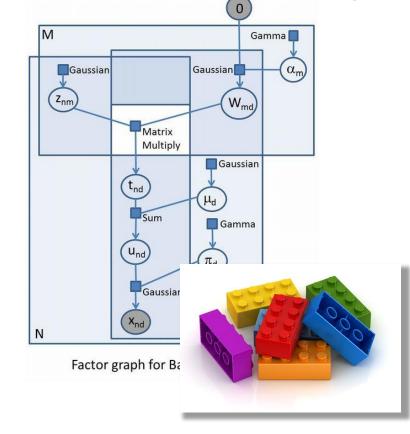
- factorize PDF
- message passing algorithms

Model-based machine learning

Choose a method



DIY - with grand unified framework



Mapping interconnections & change in capital markets

Information filtering networks to generate sparse probabilistic models (cf G-LASSO) E.g. "de-noise" a correlation matrix (Tomaso Aste & Tiziana Di Matteo)

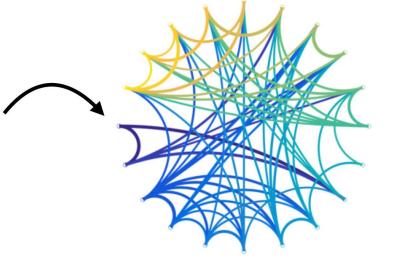








n



C BY 4.0 via Wikimedia Commons o from Madrid, Spain; CC BY 2.0 via Wikimedia Comm

edia Commons

Parsimonious interconnectedness models

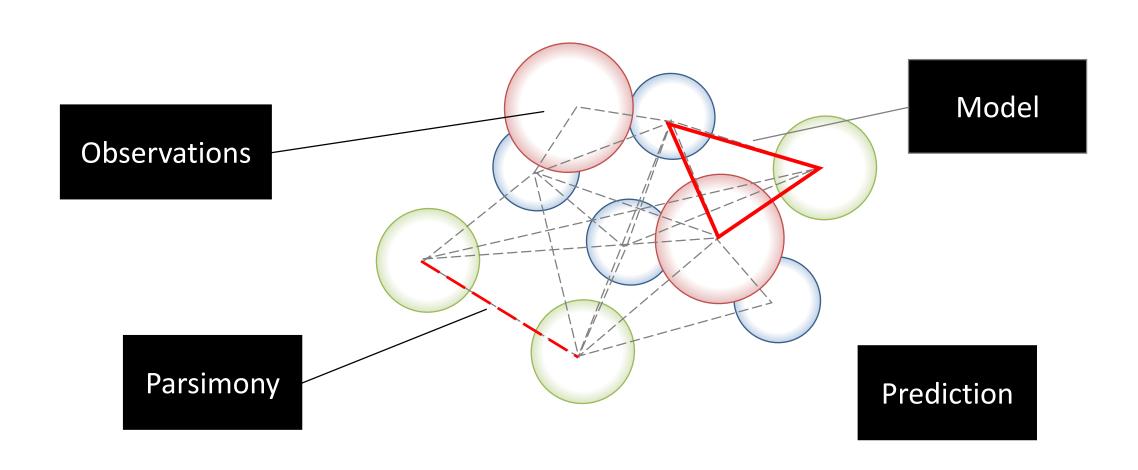
Problem

- CMSA requires quant measures for interconnectedness of products (assess, visualize)
- Desire model = MV PDF
- Public market data series: short & numerous
- Many connections in high dim'l space
- Sparse structure learning can find meaningful and parsimonious models, but slowly:
 - Constraint, score, regression based (ridge, lasso, elastic-net), decomposable models
 - Cf graphical least absolute shrinkage & selection operator (G-LASSO) (L1 reglzn)

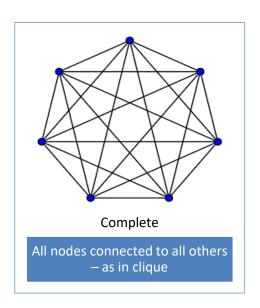
Solution

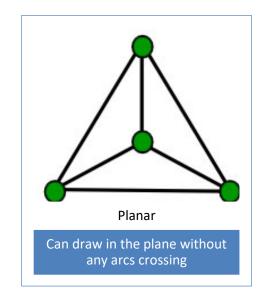
- Information Filtering Networks find maximally filtered graph, flavours:
 - Max spanning tree (MST)
 - Planar PMFG
 - Triangulated TMFG (efficient & decomposable)
 - LoGo = decomposable IFN + GMRF
- Trick: data are points
 - Don't: start with full pdf, & discard (conditionally) independent variables
 - Do: start with "similar" vars & build geometric structures (clique forest) ("knowledge")
- E.g. estimate inverse covariance of high-dim'l, noisy, short time-series
- Simplify the system of connections

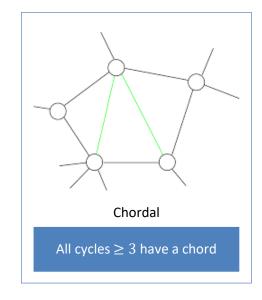
Predictive modelling for a complex world

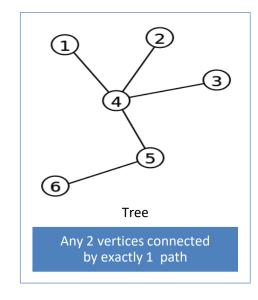


Graphs: complete, planar, chordal, tree









Predictive modelling

• Predictive modelling

Conditional probability
$$\widetilde{\mathbb{P}(B|A)} = \frac{\widetilde{\mathbb{P}(A,B)}}{\mathbb{P}(A)}$$

Conditional expectation

$$\mathbb{E}(Y|X=x) \qquad \text{(Discrete RV.)}$$

$$= \sum_{y} y \, \mathbb{P}(Y=y|X=x)$$

Conditional entropy

$$\mathbb{H}(Y|X) = -\sum_{x,y} p(x,y) \ln p(y|x)$$

 Reduction of uncertainty on Y, given past X⁻, discounting for past Y⁻, is:

$$\mathbb{H}(Y|Y^{-}) - \mathbb{H}(Y|X^{-}, Y^{-})$$
$$= \mathbb{TE}(X \to Y)$$

Transfer entropy, for linear (MV Gsn) is Granger causality

Conditional dependency

- Compare
 - Independent $p(X,Y|\bar{Z}) = p(X|\bar{Z}) \times p(Y|\bar{Z})$
 - Dependent $p(X,Y|\bar{Z}) \neq p(X|\bar{Z}) \times p(Y|\bar{Z})$
 - $\bar{Z} = Z \setminus \{X, Y\}$
- Conditional dependency
 - As hard as entire joint PDF ☺
 - Combinatoric complexity

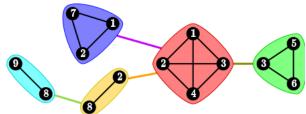
- Get inference structure from information filtering network
- Steps
 - Connect nodes that are close
 - Euclidean correlated
 - Hyperbolic mutual info $(\ln \rho)$
 - Maintain chordal property
 - Constrain clique size, planarity, info criteria
- Fast! O(N); parallelizable $O(\ln N)$

T. Aste, T. Di Matteo and S. T. Hyde, Complex networks on hyperbolic surfaces Physica A 346 (2005) 20-26.

M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, "A tool for filtering information in complex systems" Proceedings of the National Academy of Sciences of the United States of America 102, 10421 (2005).

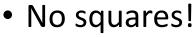
Chordal graphs





 Chordal* graphs have a recursive decomposition by clique separators into smaller subgraphs

*Or decomposable or triangulated



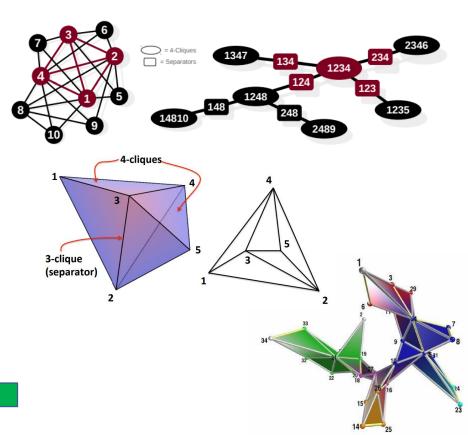




 Decomposable graphs can be parametrized by (ratios of) marginals

$$p(X) = \frac{\prod_{c} p(X_{c})}{\prod_{S} p(X_{S})^{(k_{S}-1)}}$$
Cliques (N)
Separators (N - 1)

 Estimate joint pdf of system (numerous) from pdfs of cliques and separators (few)



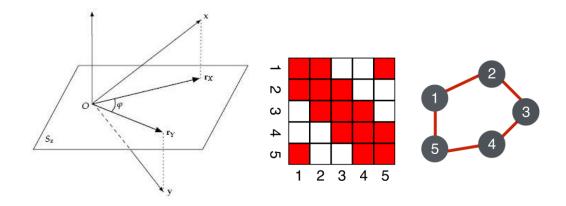
Linear / Gaussian models

Multivariate normal:

$$p(X) = \frac{1}{Z} \exp(-\sum_{i,j} X_i.J_{ij}.X_j)$$

- Zero uncertain interactions iff RVs conditionally indep
- Precision matrix (PM) $J = \Sigma^{-1}$ is sparse, w' IFN structure

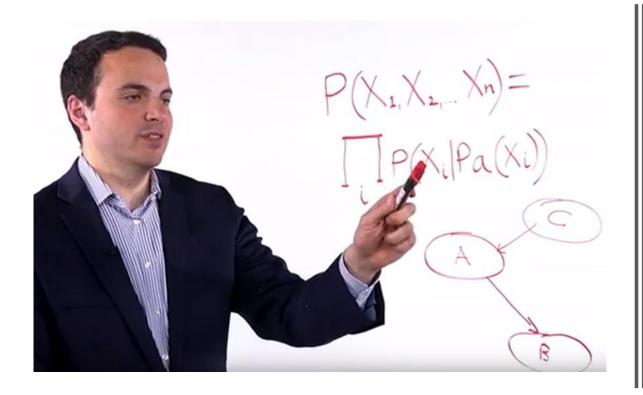
• Sparse PM from *local* inversion of cov matrix over clique forest $J_{ij} = \sum_{C} \sum_{C}^{-1} - \sum_{S} (\mathbf{k}_{S} - 1) \sum_{S}^{-1}$

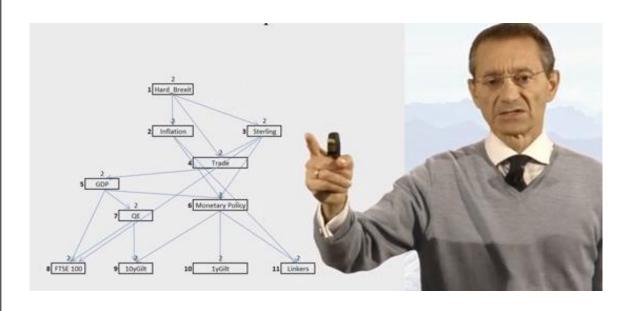


Further reading

Information filtering

- Barfuss, Wolfram, Guido Previde Massara, T. Di Matteo, and Tomaso Aste.
 2016. "Parsimonious Modeling with Information Filtering Networks." *Physical Review E* 94 (6). https://doi.org/10.1103/PhysRevE.94.062306.
- https://vimeo.com/185463870 Video
- http://bigdatafinance.eu/wp/wp-content/uploads/2017/07/Predictive-modelling-with-information-filtering-networks.pdf





Coherent stress testing

"Stress testing is in a statistical purgatory. We have some loss numbers, but who is to say whether we should be concerned about them?" Berkowitz's (1999)

Scary stress numbers – should we worry?

Stress testing with Bayesian nets

Problems

- Assess resilience of the financial system
- ST incoherent story
- Inputs
 - Scenarios
 - Macroeconomic aggregates
 - Severe. Likely?
 - Backward-looking models historical data
 - Domain knowledge (incl. causal)
- Complex
- Outputs
 - Interpret & understand
 - Aggregate diversification / amplification

Solution

- BN = intuitive, transparent structure
 - Visual; understandable
 - BN = model (assumptions)
 - Parsimonous factorization of PDF
- Aggregate sources of info
 - historical, market prices, expert opinions etc
- Updated with new info
- Rigorous
- Holistic
 - credit, market, liquidity risks
- ~ Structural equation models (SEMs)

Risk propagation – BN variables

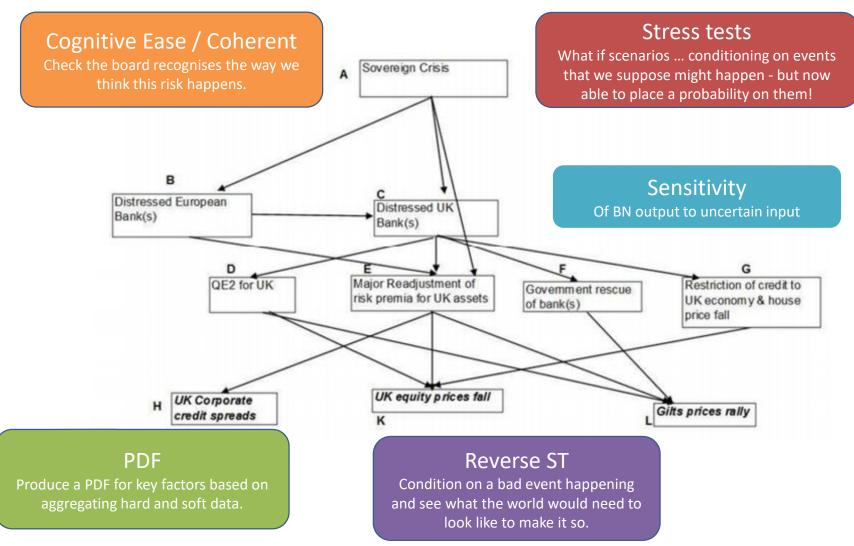
Root events = triggers e.g. break up of Euro

Connecting nodes = transmission channels e.g. failure of FI, actions of CB

Macro-financial variables e.g. GDP growth, inflation, market sentiment

Representative market indices e.g. bond yields, oil, TSX index

Granular market indices e.g. non-BM bond maturities, stock prices etc.



Further reading

Rebonato

- Rebonato, Riccardo. "How to Escape the Statistical Purgatory of Stress Testing," 2015, 26. 2015risksummit-risktesting-slides-rebonato.pdf
- ——. Riccardo Rebonato: How to Escape the Statistical Purgatory of Stress Testing, 2015. https://vimeo.com/133662430.
- Rebonato, Riccardo. Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress. Chichester, West Sussex, UK: Hoboken, NJ: Wiley, 2010.
- Rebonato. *Bayesian Nets for Stress Testing*. Accessed November 5, 2018. https://www.youtube.com/watch?time_continue=2&v=BqlImtoY0i0.
- ———. Prof. Riccardo Rebonato on "Coherent Stress Testing Should We Worry? What Should We Do?," 2016. https://www.youtube.com/watch?v=YcbqCFnl4yY.
- Rebonato, Riccardo. "A Bayesian Approach to Stress Testing and Scenario Analysis."
 ResearchGate, 2010.
 https://www.researchgate.net/publication/228232370_A_Bayesian_Approach_to_Stress Testing and Scenario Analysis.
- ———. "A Financially Justifiable and Practically Implementable Approach to Coherent Stress Testing," 2018.
- ——. "The ERI Stress Testing Tool: A Coherent Approach to Stress Testing," September 1, 2017. https://risk.edhec.edu/sites/risk/files/images/stress testing.pdf.

Deney

- Denev, Alexander. "Financial Networks as Directed Cyclic Graphs." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, February 16, 2014. https://papers.ssrn.com/abstract=2392756.
- ———. "Financial Networks as Probabilistic Graphical Models (PGM)." Cambridge, 2015.
 http://www.jbs.cam.ac.uk/fileadmin/user_upload/research/centres/risk/downloads/150909_slides_denev.pdf.
- ———. "Probabilistic Graphical Models: A New Way of Thinking in Financial Modelling." Risk.net, July 1, 2015. http://www.risk.net/node/2426700.
- ———. Probabilistic Graphical Models: A New Way of Thinking in Financial Modelling. London: Risk Books, 2015.
- ———. "Stress Testing A New Approach and Applications." Accessed February 6, 2017.

 http://www.hpcfinance.eu/sites/www.hpcfinance.eu/files/20140213_1445_Alexan_der%20Denev%20-%20RiskDynamics.pdf.
- Rebonato, Riccardo, and Alexander Denev. Portfolio Management under Stress: A Bayesian-Net Approach to Coherent Asset Allocation. Cambridge: Cambridge University Press, 2014.
- Rebonato, Riccardo, and Alexander Denev. "Coherent Asset Allocation and Diversification in the Presence of Stress Events." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 27, 2011. https://papers.ssrn.com/abstract=1824207.

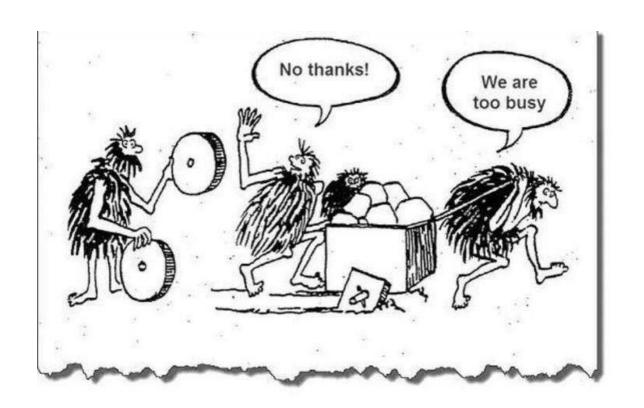
Causal inference

... I see no greater impediment to scientific progress than the prevailing practice of focusing all our mathematical resources on probabilistic and statistical inferences while leaving causal considerations to the mercy of intuition and good judgment.

-- Pearl, 1999

To know, is to know the final cause.
-- Aristotle

When I was growing up, we didn't have DAGs because we was too PO
-- Tim Simmons

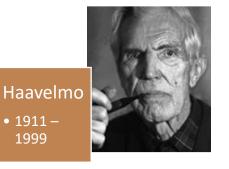


Innovators – causal inference

Wright • 1889-1988







1999

Duncan







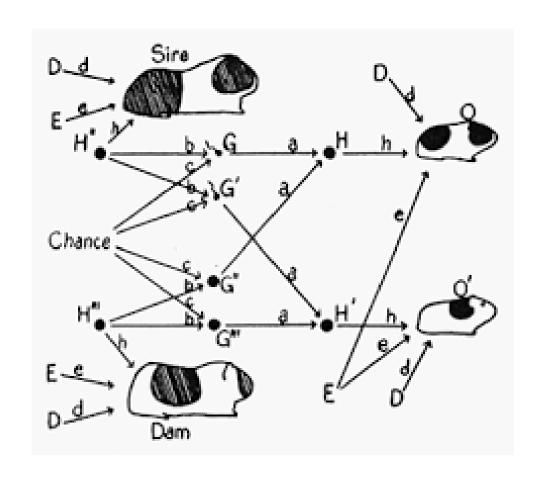
Heckman • 1944-



Causal inference – science of cause & effect

A causal model

- predicts behaviour of a system.
 - probability, of **counterfactual** claims about the system;
 - effects of interventions
 - probabilistic dependence or independence of variables included in the model.
- inverse of these inferences:
 - if we have observed
 - probabilistic correlations among variables
 - outcomes of experimental interventions
 - determine which causal models are consistent with these observations
- Wright
 - path analysis 1921
 - (Unblocked) paths are a channels along which information (correlation) can flow, and so we add across channels



Counterfactual question:
If X would have been X',
what would be the value
of Y?

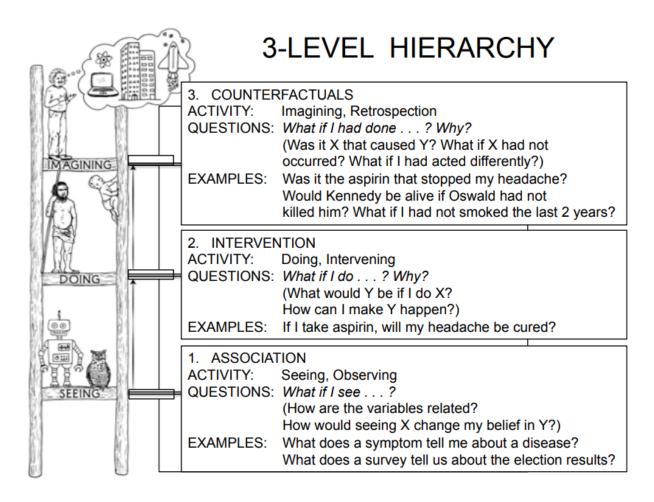
Interventionist question: If X is changed to X', what will be the value of Y?

• Experiments, Reinforcement learning, Contextual bandits.

Correlation question: How well can X predict Y?

Machine learning, Statistical estimation.

See. Do. Imagine.



Source: http://causality.cs.ucla.edu/blog/wp-content/uploads/2017/12/nips-dec2017-bw.pdf

Estimating the effect of policies on outcomes

- However, in many cases, [randomized] experiments remain difficult or impossible to implement, for financial, political, or ethical reasons, Thus, a large share of the empirical work in economics about policy questions relies on observational data—that is, data where policies were determined in a way other than through random assignment. Drawing inferences about the causal effect of a policy from observational data is quite challenging
 - Athey & Imbens.

- We then briefly discuss some new developments in the machine learning literature, which focus on the combination of predictive methods and causal questions. We argue that machine learning methods hold great promise for improving the credibility of policy evaluation, ... Overall, this article focuses on recent developments in econometrics that may be useful for researchers interested in **estimating the effect of** policies on outcomes.
 - Athey & Imbens

Further reading

- FormalisedThinking. 2010. "Reasoning with Causality: Sequence Index." Formalised Thinking (blog). August 20, 2010. https://formalisedthinking.wordpress.com/2010/08/20/pearls-formalisation-of-causality-sequence-index/.
- Huszár, Ferenc. 2018. "ML beyond Curve Fitting: An Intro to Causal Inference and Do-Calculus." InFERENCe (blog). May 24, 2018. https://www.inference.vc/untitled/.
- Kalisch, Markus, Alain Hauser, Martin Maechler, Diego Colombo, Doris Entner, Patrik Hoyer, Antti Hyttinen, et al. 2018. *Pcalg: Methods for Graphical Models and Causal Inference* (version 2.6-0). https://cran.r-project.org/package=pcalg.
- Nielsen. 2012. "If Correlation Doesn't Imply Causation, Then What Does? | DDI." 2012. http://www.michaelnielsen.org/ddi/if-correlation-doesnt-imply-causation-then-what-does/.
- Pearl, Judea. 2000. Causality: Models, Reasoning, and Inference. 1 edition. Cambridge, U.K.; New York: Cambridge University Press.
- ——. 2013. "A Probabilistic Calculus of Actions." ArXiv:1302.6835 [Cs], February. http://arxiv.org/abs/1302.6835.
- Pearl, Judea, and Dana Mackenzie. 2018. The Book of Why: The New Science of Cause and Effect. 1 edition. New York: Basic Books.
- Tikka, Santtu. 2018. Causaleffect: Deriving Expressions of Joint Interventional Distributions and Transport Formulas in Causal Models (version 1.3.9). https://cran.r-project.org/package=causaleffect.
- Tyrväinen, Sanna. 2017. "Introduction to Causal Calculus," 17.
- Winn, John. 2012. "Causality with Gates," April. https://www.microsoft.com/en-us/research/publication/causality-with-gates/.

Software

Software for building probabilistic graphical and or causal models

Software

probabilistic programming & graphical models

•Python library for working with Probabilistic Graphical Models.

•R package for learning the graphical structure of Bayesian networks, estimate their parameters and perform some useful inference.

•Framework for running Bayesian inference in GMs

- •It can also be used for probabilistic programming.
- Recently open-sourced by MSFT
- Script with F#

 Probabilistic programming language for statistical inference written in C++.

•Script with PyStan & Rstan.

Prob prog package for Python

*Commercial

- Fit Bayesian models: Markov chain Monte Carlo (MCMC) and variational inference (VI).
- •Used by Quantopian.
- •Built using Theano

Pgmpy



BNLearn



Infer.NET



Stan



PyMC3



- Python library for probabilistic modeling & inference,
- Built on **TensorFlow**.

Gaussian graphical models in scikit-learn

 Python library to estimate causal effects.; based on a unified language for causal inference, combining causal graphical models and potential outcomes frameworks. Bayesia develops AI software for ML, knowledge modeling, analytics, simulation and optimization. • Avis Nigra offers expertise in scenario-based risk management by constructing relevant and tractable models of risks that are not accessible by statistics alone.

Edward



Skggm



DoWhy



BayesiaLab*



Avis Nigra*



Conclusions

Probabilistic graphical models as a tool for

- sparse interconnectedness models from historical market data and
- coherent stress tests for financial system resilience assessment

Conclusions

Parsimonious interconnectivity

- Aste, Di Matteo et al
- Information filtering networks
- Graphical representation of strongest dependencies e.g. precision matrix J
- Eliminating weak dependences as hard as full precision matrix ☺
- Construct quickly, sparse matrix as clique forest
- Chordal graph pdf factorizes ©
- Sparse regression prediction
- Deduce causal relationships (output)

Coherent stress tests

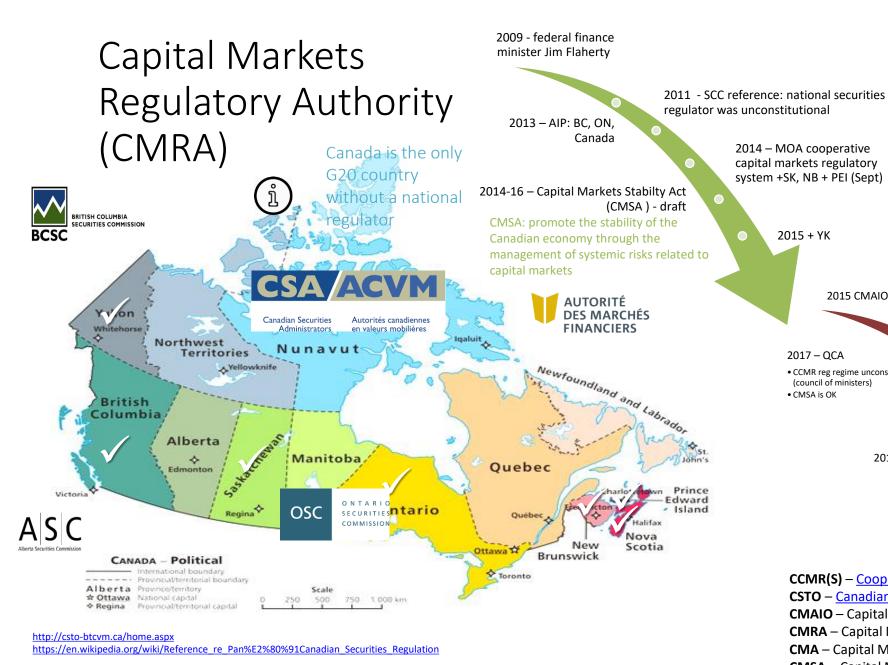
- Rebonato, Denev
- Should we worry?
- Causation central (input)
- Framework for quant stories
- Combine frequentist & subjective uncertainty
- Intuitive, understandable (challengeable) by executives
- Sensitivity analysis which inputs
- Produce pdfs for market indices driving financial system

The End

Thank you for listening

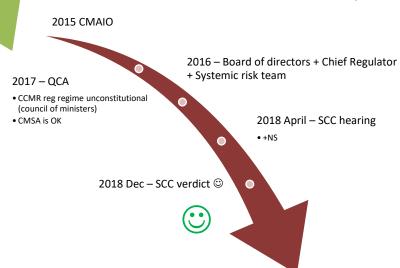
Appendix: Canadian context

History, geography of the Canadian regulatory landscape



"The preservation of capital markets and the maintenance of Canada's financial stability ... do not justify a wholesale takeover of the regulation of the securities industry"

2020 - Launch



CCMR(S) - Cooperative Capital Markets Regulatory (System) (2014)

CSTO – Canadian Securities Transition Office (2009)

CMAIO - Capital Markets Authority Implementation Organization (2015)

CMRA – Capital Markets Regulatory Authority (2020)

CMA - Capital Markets Act ()

2015 + YK

CMSA - Capital Markets Stability Act (2014-16)

CANADIAN REGULATORY LANDSCAPE

