

Gene Regulatory Network Inference

Introduction to GRN Inference

Sergio Peignier

sergio.peignier@insa-lyon.fr

Associate Professor

INSA Lyon

Biosciences department

Table of contents

1. Principles

2. Main families

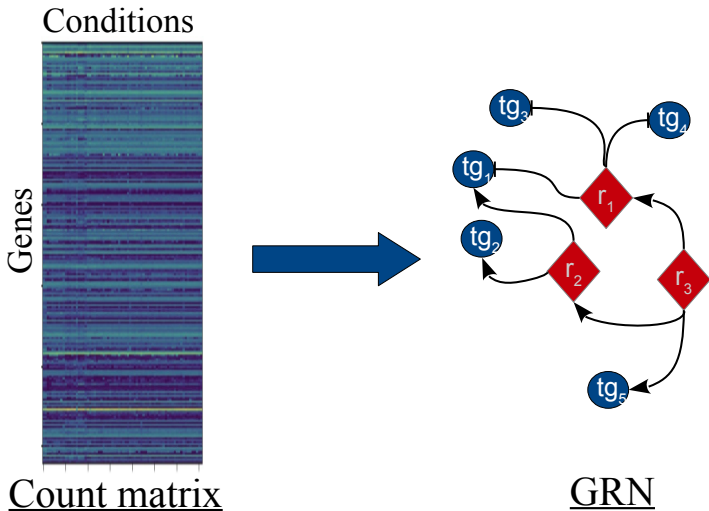
Principles

Input Data

- **Gene expression matrix**
(RNAseq or Microarray)
- n Rows : **Genes**
($n \sim 10^4$)
- p Columns : **Conditions**
($n \sim 10^1 \rightarrow 10^2$)
- Here **TFs** and **TGs** only



How to infer GRNs from gene expression data?



Which TFs regulate the TG?

- X_{tf_1} , X_{tf_2} and X_{tf_3} : **Gene expressions** of 3 TFs
- X_{tg_0} : **Gene expression** of a TG

$$X_{tf_1} = \begin{pmatrix} 1.2 \\ 0.0 \\ 3.3 \\ 0.0 \end{pmatrix} \quad X_{tf_2} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.5 \\ 1.5 \end{pmatrix} \quad X_{tf_3} = \begin{pmatrix} 0.0 \\ 2.1 \\ 0.3 \\ 4.1 \end{pmatrix} \quad X_{tg_0} = \begin{pmatrix} 0.8 \\ 0.1 \\ 1.4 \\ 0.0 \end{pmatrix}$$

$$tf_1 \stackrel{?}{\rightarrow} tg_0 \quad tf_2 \stackrel{?}{\rightarrow} tg_0 \quad tf_3 \stackrel{?}{\rightarrow} tg_0$$

Intuitive answer

- X_{tf_1} , X_{tf_2} and X_{tf_3} : **Gene expressions of 3 TFs**
- X_{tg_0} : **Gene expression of a TG**

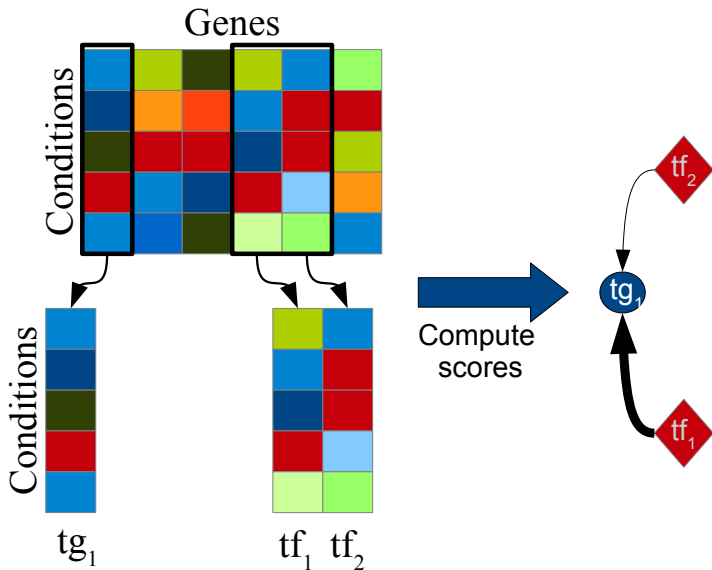
$$X_{tf_1} = \begin{pmatrix} 1.2 \\ 0.0 \\ 3.3 \\ 0.0 \end{pmatrix} \quad X_{tf_2} = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.5 \\ 1.5 \end{pmatrix} \quad X_{tf_3} = \begin{pmatrix} 0.0 \\ 2.1 \\ 0.3 \\ 4.1 \end{pmatrix} \quad X_{tg_0} = \begin{pmatrix} 0.8 \\ 0.1 \\ 1.4 \\ 0.0 \end{pmatrix}$$

$$\boxed{tf_1 \rightarrow tg_0}$$

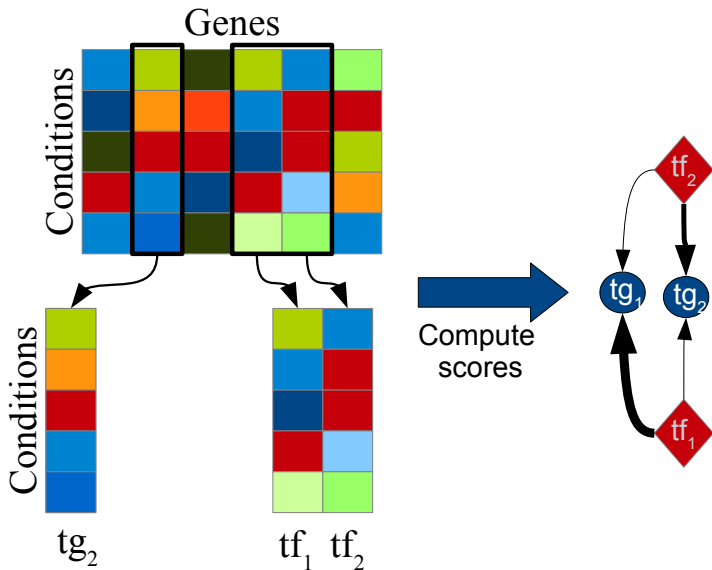
$$tf_2 \not\rightarrow tg_0$$

$$\boxed{tf_3 \rightarrow tg_0}$$

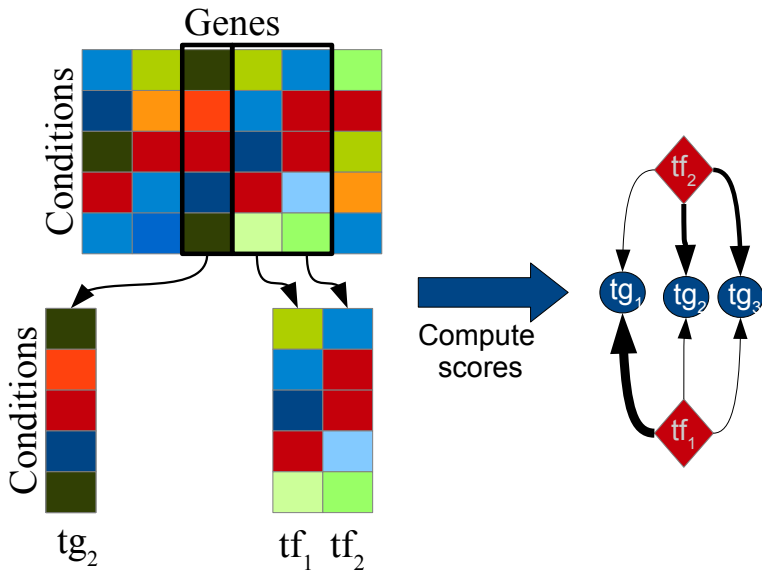
Score possible relationships



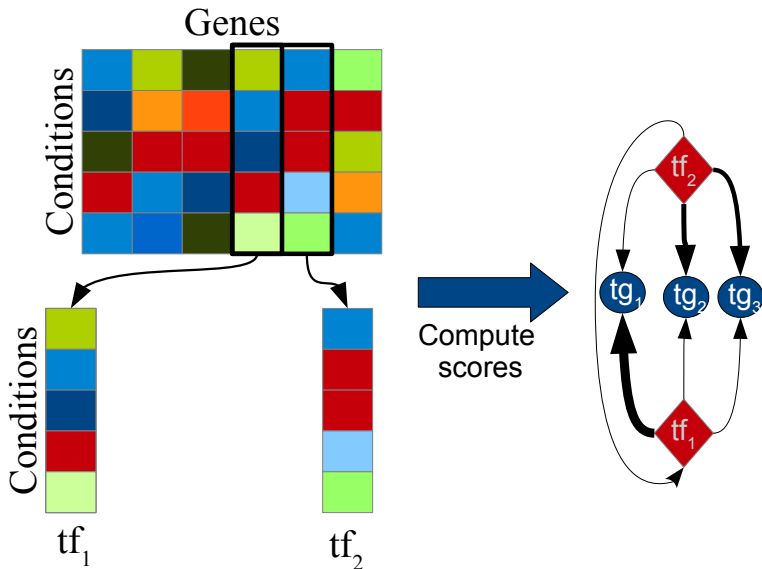
Score possible relationships



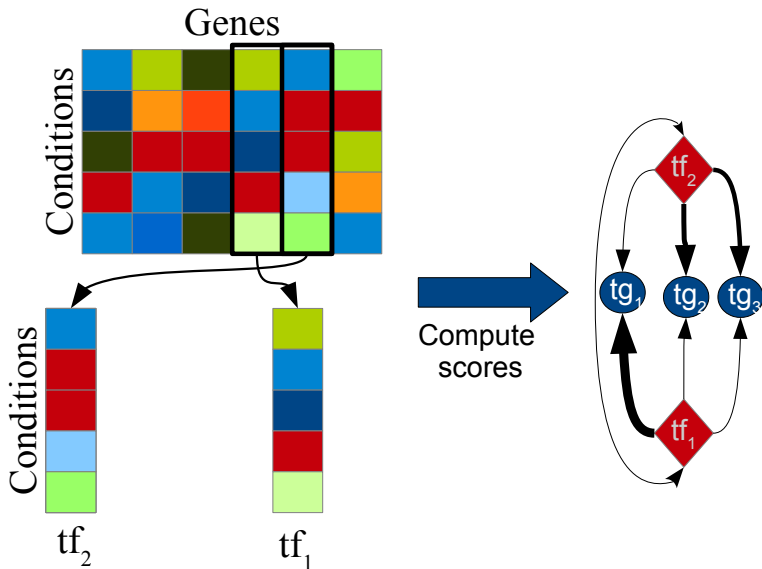
Score possible relationships



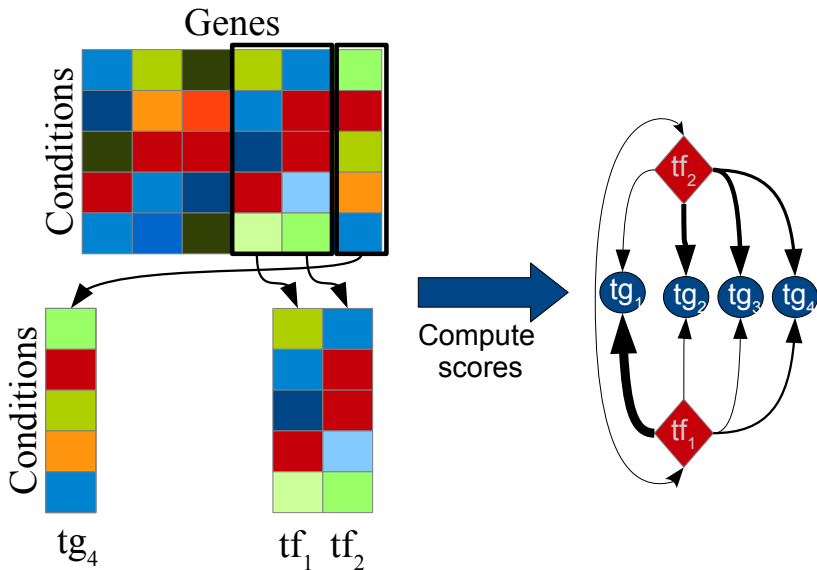
Score possible relationships



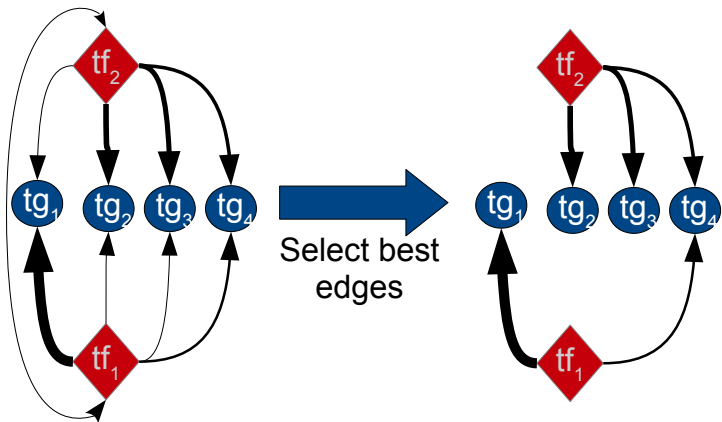
Score possible relationships



Score possible relationships



Score possible relationships



Main families

Main GRN Inference Families

Methodologically the main GRN Inference techniques are based on:

- Correlation
- Mutual Information
- Regression
- Bayesian inference
- Model based

Main GRN Inference Families

Methodologically the main **GRN Inference** techniques are based on:

- **Correlation:** The **score** computation relies on **correlation measures**
- **Mutual Information**
- **Regression**
- **Bayesian inference**
- **Model based**

Main GRN Inference Families

Methodologically the main GRN Inference techniques are based on:

- Correlation
- Mutual Information: The score computation relies on Mutual Information measures
- Regression
- Bayesian inference
- Model based

Main GRN Inference Families

Methodologically the main **GRN Inference** techniques are based on:

- Correlation
- Mutual Information
- Regression: The **score** is computed using a regression algorithm (e.g., linear regression)
- Bayesian inference
- Model based

Main GRN Inference Families

Methodologically the main **GRN Inference** techniques are based on:

- Correlation
- Mutual Information
- Regression
- **Bayesian inference**: The objective is to **fit** a **Bayesian network**, i.e., a **probabilistic model** of the GRN
- Model based

Main GRN Inference Families

Methodologically the main **GRN Inference** techniques are based on:

- Correlation
- Mutual Information
- Regression
- Bayesian inference
- **Model based**: The objective is to **fit** an **ODE** or a **Boolean Network model** of the **GRN**, using **optimization techniques** (e.g., genetic algorithms, simulated annealing)

Practical session objective

- **Explore and implement GRN Inference methods** based on:
 - Correlation
 - Mutual Information
 - Regression
 - Classification
- **Apply** them to **real world data**
- **Compare their results**