

ONLINE LEARNING AND CAUSALITY

THIS IS A TEMPORARY TITLE PAGE
It will be replaced for the final print by a version
provided by the registrar's office.

Thèse n. 1234 2020
présentée le 6 juillet 2020
à la Faculté des sciences de base
laboratoire SuperScience
programme doctoral en SuperScience
École polytechnique fédérale de Lausanne
pour l'obtention du grade de Docteur ès Sciences
par

Paolino Paperino

acceptée sur proposition du jury :

Prof Name Surname, président du jury
Prof Name Surname, directeur de thèse
Prof Name Surname, rapporteur
Prof Name Surname, rapporteur
Prof Name Surname, rapporteur

Lausanne, EPFL, 2020



Wings are a constraint that makes
it possible to fly.
— Robert Bringhurst

To my parents...

Acknowledgements

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Lausanne, July 6, 2020

D. K.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Contents

Acknowledgements	i
Abstract	iii
Introduction	1
I Body 1 (Can change this latter)	3
1 Tables and Figures	5
1.1 Tables	5
1.2 Figures	6
II Body 2	11
2 Mathematics	13
2.1 Very important formulas	13
3 Causal Discovery	15
3.1 Causal algo	15
3.2 Proof of consistency: A tale of two bounds	15
3.2.1 Bounding the false false postive	17
3.2.2 Bounding the false negatives	17
A An appendix	19
Bibliography	21
Bibliography	21

Introduction

The endeavour of science is in some sense the uncovering of causal structures: does the mass of an object influence its acceleration in free fall? What genomes influence height (etc)? The ability to do experiments in Physics is what has allowed to confirm or uncover relations among objects; indeed, this is also how we learn best, by tweaking a system and having a direct feedback which allows us to evaluate our mental models. In the realm of causality, such a setting is what is known as the "Causal intervention framework" (need to check). Give quickly example about smoking, and illustrate how it would be harder without interventions.

An excellent question about a multivariable causal system is to ask, "what is the minimum number of interventions one has to do to achieve some alpha-confidence about the causal effect"

How can we tackle causality?

In the absence of noise, and the process is bijective, then it is impossible to distinguish, if however, ...

Shannon answered the question: given the most simple communication system: "How reliably can we communicate given a certain noise level"

In some sense what we would like to answer is, given a certain noise level, how reliably can we predict the causal relation.

Some points:

1. In causality we use noise, whereas in virtually all other domains such as communication theory the aim is combat noise.

Interestingly yet again, the Gaussian case ends up being a difficulty case. For instance, the motivation to look at the AGN additive gaussian noise channel is that the gaussian is the most difficult distribution in the entropic sense; but so it is as well in the binary case setting due to:

thm.

A non-numbered chapter...

Introduction

Talk about SNR, role with shannon, and how it affects prediction in a reverse way here! Cite shanon!

Body 1 (Can change this latter) Part I

1 Tables and Figures

In this chapter we will see some examples of tables and figures.

1.1 Tables

Let's see how to make a well designed table.

The table 1.1 is a floating table and was obtained with the following code:

```
1 \begin{table}[tb]
2 \caption[A floating table]{A floating table.}
3 \label{tab:example}
4 \centering
5 \begin{tabular}{ccc}
6 \toprule
7     name      & weight & food  \\
8 \midrule
9     mouse    & 10 g   & cheese \\
10    cat      & 1 kg   & mice  \\
11    dog      & 10 kg  & cats  \\
12    t-rex    & 10 Mg  & dogs  \\
13 \bottomrule
14 \end{tabular}
15 \end{table}
```

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget,

Table 1.1 – A floating table.

name	weight	food
mouse	10 g	cheese
cat	1 kg	mice
dog	10 kg	cats
t-rex	10 Mg	dogs

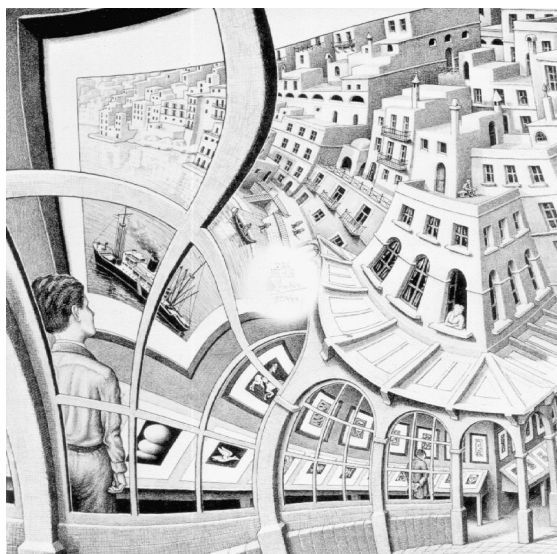


Figure 1.1 – A floating figure (the lithograph *Galleria di stampe*, of M. Escher, got from <http://www.mcescher.com/>).

consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

1.2 Figures

Let's see now how to put one or several images in your text.

The figure 1.1 is a floating figure and was obtained with the following code:

```
1 \begin{figure}[tb]
2 \centering
3 \includegraphics[width=0.5\columnwidth]{galleria_stampe}
4 \caption[A floating figure]{A floating figure ... }
5 \label{fig:galleria}
```

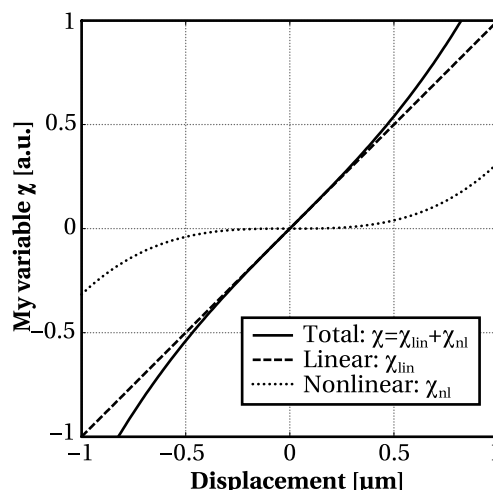



Figure 1.2 – A floating figure with text typeset in "Utopia Latex", a font provided in the template folder for typesetting figures with greek characters. The text has been "outlined" for best compatibility with the repro during the printing.

6 \end{figure}

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

The figure 1.3 is a floating figure and was obtained with the following code:

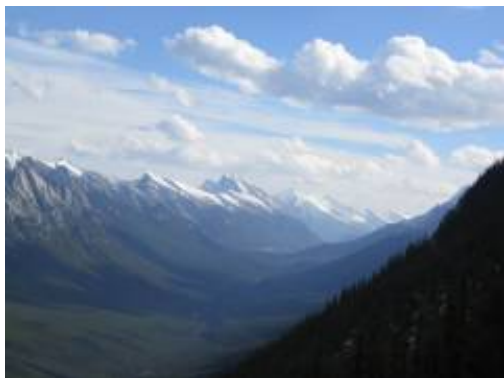
```
1 \begin{figure}[tb]
2 \centering
3 \subfloat[Asia personas duo.]
4 {\includegraphics[width=.45\columnwidth]{lorem}} \quad
5 \subfloat[Pan ma signo.]
6 {\label{fig:ipsum}%
7 \includegraphics[width=.45\columnwidth]{ipsum}} \\\
```



(a) Asia personas duo.



(b) Pan ma signo.



(c) Methodicamente o uno.



(d) Titulo debitas.

Figure 1.3 – Tu duo titulo debitas latente.

```

8 \subfloat[Methodicamente o uno.]
9 {\includegraphics[width=.45\columnwidth]{dolor}} \quad
10 \subfloat[Titulo debitas.]
11 {\includegraphics[width=.45\columnwidth]{sit}}
12 \caption[Tu duo titulo debitas latente]{Tu duo titulo debitas latente.}
13 \label{fig:esempio}
14 \end{figure}

```

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst.

Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Body 2 Part II

2 Mathematics

In this chapter we will see some examples of mathematics.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.1 Very important formulas

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

$$\frac{d}{dt} \begin{bmatrix} P_0 \\ P_I \\ P_T \end{bmatrix} = \begin{bmatrix} \frac{P_I}{\tau_{10}} + \frac{P_T}{\tau_T} - \frac{P_0}{\tau_{ex}} \\ -\frac{P_I}{\tau_{10}} - \frac{P_I}{\tau_{isc}} + \frac{P_0}{\tau_{ex}} \\ \frac{P_I}{\tau_{isc}} - \frac{P_T}{\tau_T} \end{bmatrix} \quad (2.1)$$

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus

adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

$$\bar{I}_f(\vec{r}) = \gamma(\vec{r}) \left(1 - \frac{\tau_T P_T^{eq} \left(1 - \exp\left(-\frac{(T_p - t_p)}{\tau_T}\right) \right)}{1 - \exp\left(-\frac{(T_p - t_p)}{\tau_T} + k_2 t_p\right)} \times \frac{(\exp(k_2 t_p) - 1)}{t_p} \right) \quad (2.2)$$

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

3 Causal Discovery

3.1 Causal algo

blup di blue

3.2 Proof of consistency: A tale of two bounds

The setup was the linear ANM:

$$\begin{cases} Y = aX + E_Y \\ X \perp\!\!\!\perp E_Y, X \sim p_x, E_Y \sim p_{E_Y} \end{cases}$$

From n samples (X_i, Y_i) we estimate \hat{f}_Y by regressing X on Y and \hat{f}_X for the reverse model. We then compute the residuals

$$\hat{e}_Y = Y - \hat{f}_Y(X) \tag{3.1}$$

$$\hat{e}_X = X - \hat{f}_X(Y) \tag{3.2}$$

We note that for the ease of analysis, it would first be wise to use some fraction of the data to first estimate the regression, and then use the remanining for the test.

For n large enough we have that

$$\hat{e}_Y \approx E_Y \sim P_{E_Y}$$

Chapter 3. Causal Discovery

The idea is then to first discretise¹ P_{E_Y} into m bins, call this discrete distribution Q . We apply the same discretization to obtain $B = (b_1, \dots, b_m)$ from \hat{e}_Y and $\tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_m)$ from \hat{e}_X .

We then decide the causal direction as follows

$$\begin{cases} X \rightarrow Y & \text{if } C \leq W \\ Y \rightarrow X & \text{if } C > W \end{cases}$$

Where

$$C = \|B - U\|_{L_1}$$

$$W = \|\tilde{B} - U\|_{L_1}$$

$$\text{s.t. } U = (\frac{1}{m}, \dots, \frac{1}{m}).$$

Given our assumption about the **ANM**, the probability to output the correct causal direction is:

$$P_{\text{correct}} = \mathbb{P}[C \leq W]$$

We next upper bound this quantity in order to show consistency

$$\mathbb{P}[C \leq W] \geq \mathbb{P}\left[\bigcup_{\tau \in \mathbb{Q}} C \leq \tau \cap W > \tau\right] \quad (3.3)$$

$$\geq \mathbb{P}[C \leq \tau \cap W > \tau] \quad (3.4)$$

$$\geq \mathbb{P}[C \leq \tau] - \mathbb{P}[W \leq \tau] \quad (3.5)$$

The first inequality is due to the fact that we are only taking the union in the rationals². The second inequality is done by looking at the probability of a fixed τ ; and the final one follows by:

$$1 \geq \mathbb{P}[C \leq \tau \cup W > \tau] = \mathbb{P}[C \leq \tau] + \mathbb{P}[W > \tau] - \mathbb{P}[C \leq \tau \cap W > \tau]$$

We will next find appropriate bounds for $\mathbb{P}[C \leq \tau]$ and $\mathbb{P}[W \leq \tau]$.

¹We do so in a naive manner we split it uniformly into m bins.

²We note that we can only take unions over countable sets; recall also that the rationals are dense in the irrationals, so the inequality is very close to equality (and in practice and among friends it would be).

3.2.1 Bounding the false false positive

We will first lower bound $\mathbb{P}[C \leq \tau]$ by upper bounding the complement event.

$$\mathbb{P}[C \geq \tau] = \mathbb{P}\left[\sum_{i=1}^m |b_i - \frac{1}{m}| \geq \tau\right] \quad (3.6)$$

$$\leq \mathbb{P}\left[m \max_i |b_i - \frac{1}{m}| \geq \tau\right] \quad (3.7)$$

$$= \mathbb{P}\left[\bigcup_i |b_i - \frac{1}{m}| \geq \frac{\tau}{m}\right] \quad (3.8)$$

$$\leq m \mathbb{P}\left[|b_0 - \frac{1}{m}| \geq \frac{\tau}{m}\right] \quad (3.9)$$

$$\leq m 2 \exp\left(-2n \frac{\tau^2}{m^2}\right) \quad (3.10)$$

The second to last inequality follows by the union bound and by noting that all b_i s are the same since they are discretized empirical distribution coming from a uniform source. For the final inequality we use Hoeffding's inequality.

3.2.2 Bounding the false negatives

Recall that what is left to bound is the following quantity, $\mathbb{P}[W \leq \tau]$; for this we first define the following set of probability distributions:

$$\Gamma_\tau = \{\pi \in \Delta_m : \|\pi - U\|_{L_1} \leq \tau\}$$

Where the Δ_m is the m dimensional simple and U the uniform vector as before.

Observe that:

$$\{W \leq \tau\} = \{\tilde{B} \in \Gamma_\tau\}$$

In essence, we are asking: "what is the chance that the realisation of \tilde{B} – which is the empirical distribution of some distribution Q – lies inside some set of distributions Γ_τ .

We note that bounding this kind of event is exactly what Sanov's theorem³ gives us, an important result from large deviation theory that also exploits concentration of measure.

Let $\mathbf{x} = (x_1, \dots, x_n)$ be a sequence of n each drawn independently from a finite universe U with

³See the section on Information Theory and statistics in Cover (1999)

Chapter 3. Causal Discovery

$|U| = m$. Denote by $P_{\mathbf{x}}$ the empirical distribution – or type – for a given sequence \mathbf{x} . Let Q^n be the product distribution n independent samples of Q .

Theorem 1 (Sanov's theorem) *Let Π be a set of distributions on U , and $m = |U|$. Let*

$$P^* = \operatorname{argmin}_{P \in \Pi} D(P \| Q)$$

Then

$$\mathbb{P}_{Q^n} [P_{\mathbf{x}} \in \Pi] \leq (n+1)^m 2^{-nD(P^* \| Q)}$$

Applying the above theorem, and noting that Γ_τ takes the place of Π , \tilde{B} that of $P_{\mathbf{x}}$ and the discretized distribution $\hat{e}_X = X - \hat{f}_X(Y)$ that of Q we get:

$$\mathbb{P} [W \leq \tau] = \mathbb{P} [\tilde{B} \in \Gamma_\tau] \leq (n+1)^m 2^{-nD(\tau)} \quad (3.11)$$

Where $D(\tau) := D(P^* \| Q)$, we make the τ relation explicit to keep in mind that the minimisation is constrained to the set Γ_τ which depends on τ .

We remark that the only place of concern is if $D(P^* \| Q) = 0$; assuming however that $Q \neq U$, then there will be some τ s.t. $Q \notin \Gamma_\tau$ and thus $D(P^* \| Q) \neq 0$.

We can now conclude by putting everything together; recall that we had shown that we could bound the success probability as follows:

$$\mathbb{P} [C \leq W] \geq \mathbb{P} [C \leq \tau] - \mathbb{P} [W \leq \tau] \quad (3.12)$$

$$\geq 1 - 2m \exp\left(-2n \frac{\tau^2}{m^2}\right) - (n+1)^m 2^{-nD(\tau)} \quad (3.13)$$

This, if we fix m , and if there exists some τ s.t. $D(\tau) > 0$ then we get consistency by letting $n \rightarrow \infty$.

We note that to get the best bound we may maximise the r.h.s. w.r.t. τ .

A An appendix

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.



Bibliography

Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.