Prep Meeting Week 36

# Important Dates

Days to go until project defense (if all works out): **45 Days: 14/07/2022.**

Days to go until the thesis submission (if all works out): **31 Days 30/06/2022.**

# Assessment Committee

Any news from the assessment committee? Nothing in my cc. or inbox.

# Thesis Writing

Wrote Chapter 1 and Chapter 2 of thesis, overall pleased with the results / storytelling / flow.

Now working on Chapter 3 and structuring of the three methods chapters. Chapter 3 will contain more the broad spectrum and concepts of structure learning, rather than any of the concrete algorithms we will be discussing. The concrete algorithms we will be discussing follow more in depth in the three methods chapters.

In Chapter 3, we will also discuss some interesting methods outside of the scope of this thesis, such as constraint-based methods like IC and FCI and PC, as well as noise structure based methods such as (VAR)LiNGAM, as I think they are interesting and sketch different approaches to learn the structure. Score-based methods, which are about our methods, permutation-based methods, continuous methods, iterative methods.

Chapter 7 needs to get some work in, but I have a good idea on how to do it. Simulated data will see acyclic matrices W, as well as cyclic matrices W, where the “best” acyclic matrix needs to be chosen. Also structural equation models. Methods that will be considered are:

* *Exhaustive is applicable.*
* *Random Walk.*
* *MCMC in the four different components.*
* *NOTEARS.*
* DAG-LASSO.
* *DAG-OMP.*
* DAG-OLS.

Italic methods are suitable for cyclic / SEM, the others are not good in those scenarios.

As well as real-life methods. There, we will simply pick only one method, such as DAG-OMP. This will showcase the interpretability and some nice investigations.

# Regret of the AR(1) with mean setting.

**Setting.** We generate data according to an AR(1) with autoregressive coefficient *a* and mean *b*, a total of *T* timesteps. We compute the (average) *expected error* of the following two models:

1. *a = a0,* and *b* ML-estimated using LOOCV given *a*, also without LOOCV.
2. *a* and *b* ML-estimates using LOOCV.

The *expected error* of using estimated *a* and *b*, whereas *a0* and *b0* were the original values is:

We verify the regret, which we define as

the average expected error of 1 / the average expected error of 2.

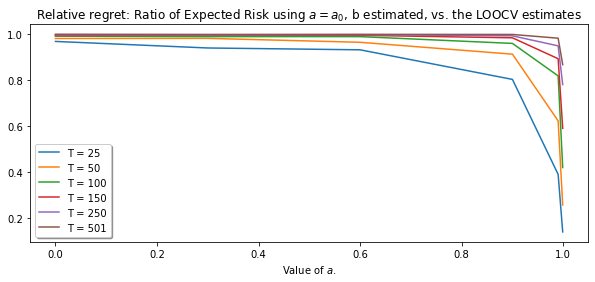
Now, we expect model 1 to always achieve the best error, as it already has a perfect estimate of *a*. For small values of *T* or values of *a* close to one, we see that the estimate for *a* in model 2 is off quite badly, and therefore the ratio drops dramatically for small *T* / *a* close to one.

So, we do not see the “bump” behavior that we expected. For small *T* / *a* close to one, the estimate of *a* is already perfect, so for example, the covariance component is equal to zero:

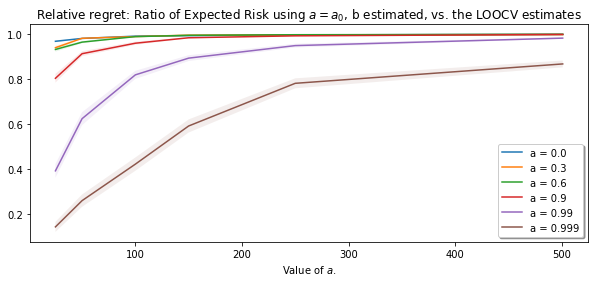
*Cov\_x \* (a0 – a)^2 = (1 / (1 – a0^2) + b0^2) \* (a0 – a)^2*.

For *a0 = 0.999* even if *a* is estimated as *a = 0.95*, the regret is already smaller than 0.45, whereas we did not account for a lot of components.

**As a function of *a*.**



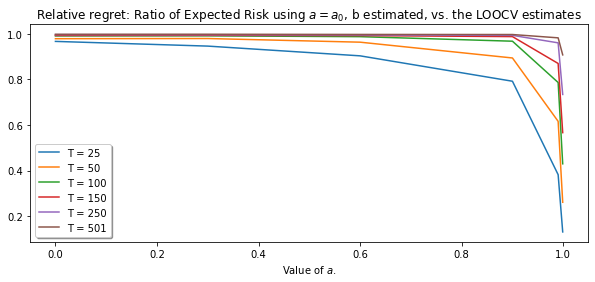
**As a function of *T.***



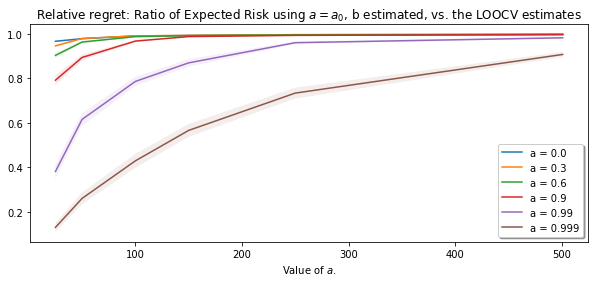
We do not see the behavior we expected. For *T* small or *a* close to one, we know that the probability of success is slightly lower. However, we hoped that this would not be bad, as the we could argue that both models would be quite bad. However, this is not the case as we see; having the correct estimate for *a* is simply too good when we look at the expected error.

**This is also the case when we do not use *LOOCV* for *b*. The plots are almost similar.**

**As a function of *a*.**

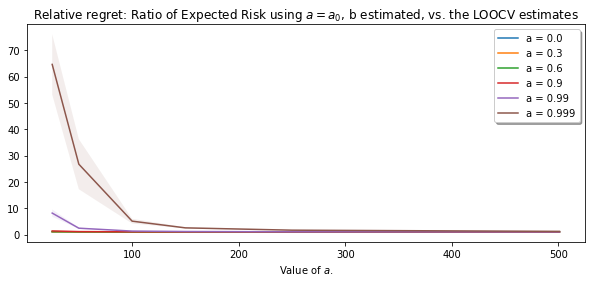


**As a function of *T*.**



**Inverse, not very useful (regret now a / b instead of b / a).**

**As a function of *T***

****

**As a function of *a*.**

****