Masters thesis topic

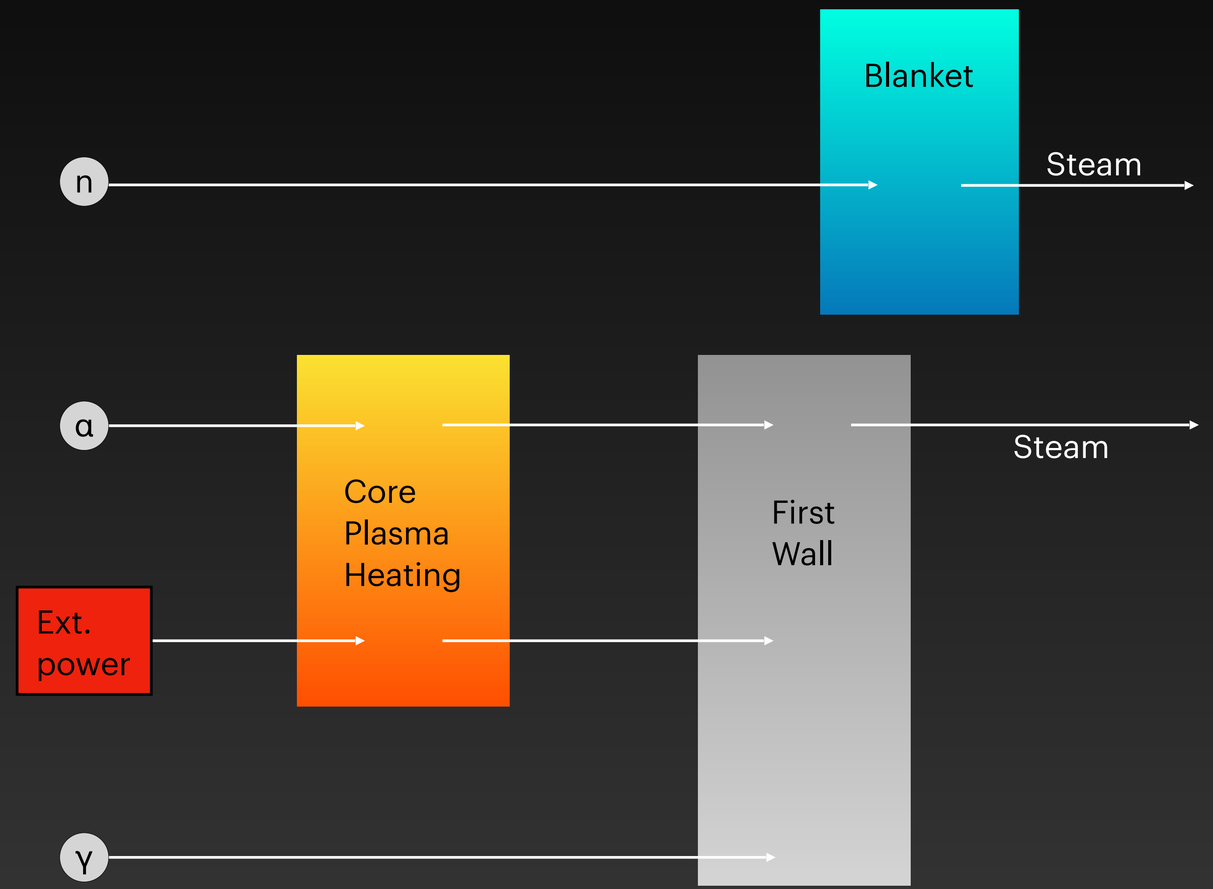
15/03/2022

Structure

- Introduction
- The problem
- Identify the cause
- Collinearity
- Classification

Introduction
Overview

- $S_\alpha + S_h = S_B + S_\kappa$
- $S_\kappa = \frac{1}{V}\int_A \mathbf{q}\,.\,\mathrm{d}\mathbf{A} = -2\frac{\kappa}{r}\frac{\partial T}{\partial r}$
- $\frac{1}{V}\int_A \mathbf{q}\cdot\mathrm{d}\mathbf{A} = \frac{3}{2}\frac{p}{\tau_E}$



The problem

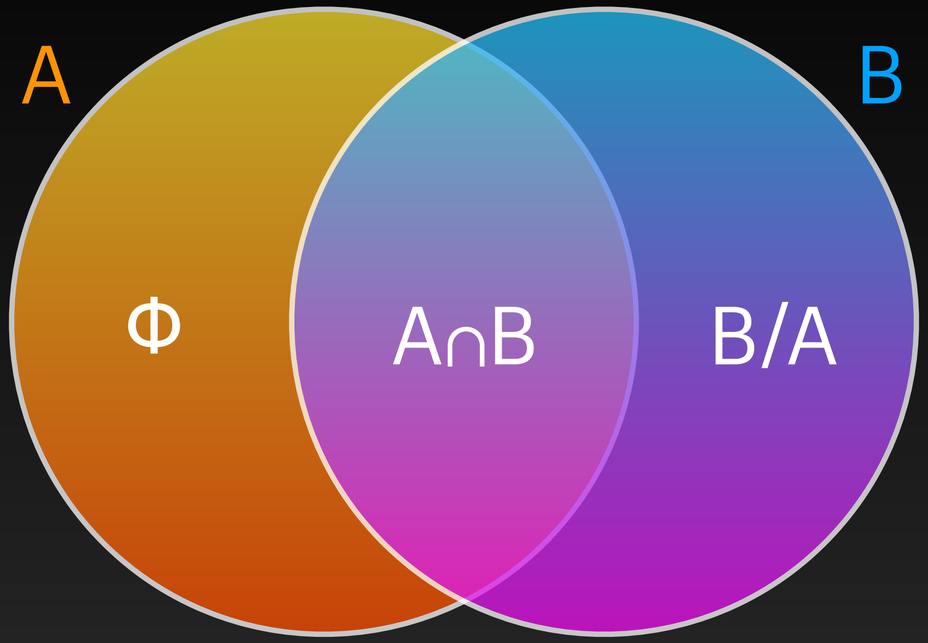
Energy confinement time

$$\tau_{\mathrm{E,th}} = \alpha_0 I_{\mathrm{p}}^{\alpha_I} B_{\mathrm{t}}^{\alpha_B} \overline{n}_{\mathrm{e}}^{\alpha_n} P_{\ell,\mathrm{th}}^{\alpha_P} R_{\mathrm{geo}}^{\alpha_R} \kappa_{\mathrm{a}}^{\alpha_\kappa} \epsilon^{\alpha_\epsilon} M_{\mathrm{eff}}^{\alpha_M}$$

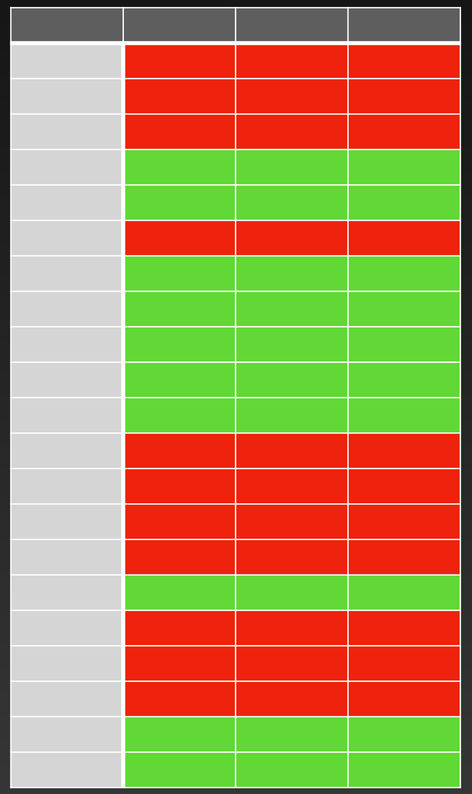
Progression of regression coefficients

Database	α0	αI	αB	αP	αn	αM	αR	αe	αk	no_of_points
DB2.8	0.0913	0.776	0.323	-0.666	0.437	0.189	2.22	0.578	0.4	1310
DB2.8 ITER	0.136	0.963	0.147	-0.664	0.334	0.3	1.84	0.372	0.6	732
DB4	0.115	0.794	0.251	-0.616	0.389	0.0646	1.97	0.712	0.471	3237
DB4 ITER	0.157	1.06	-0.0129	-0.615	0.322	0.155	1.59	0.378	0.725	2275
DB5.2.3	0.128	1.13	0.0724	-0.683	0.175	0.214	1.45	0.0636	0.28	6250
DB5.2.3 ITER	0.147	1.38	-0.182	-0.669	0.112	0.303	1.07	-0.136	0.712	5220

Identify the cause



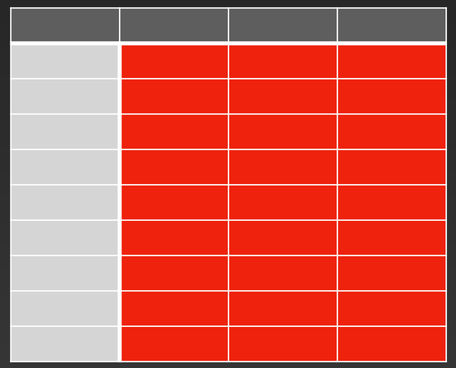
B



αR = 1.07

=

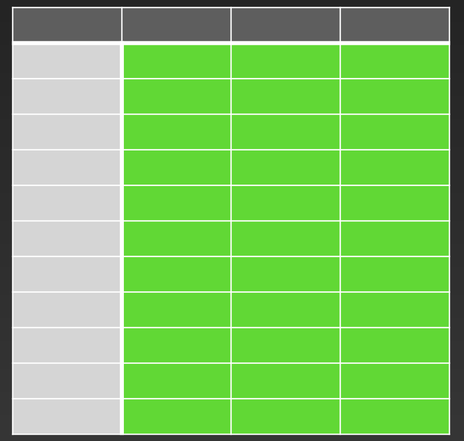
A∩B

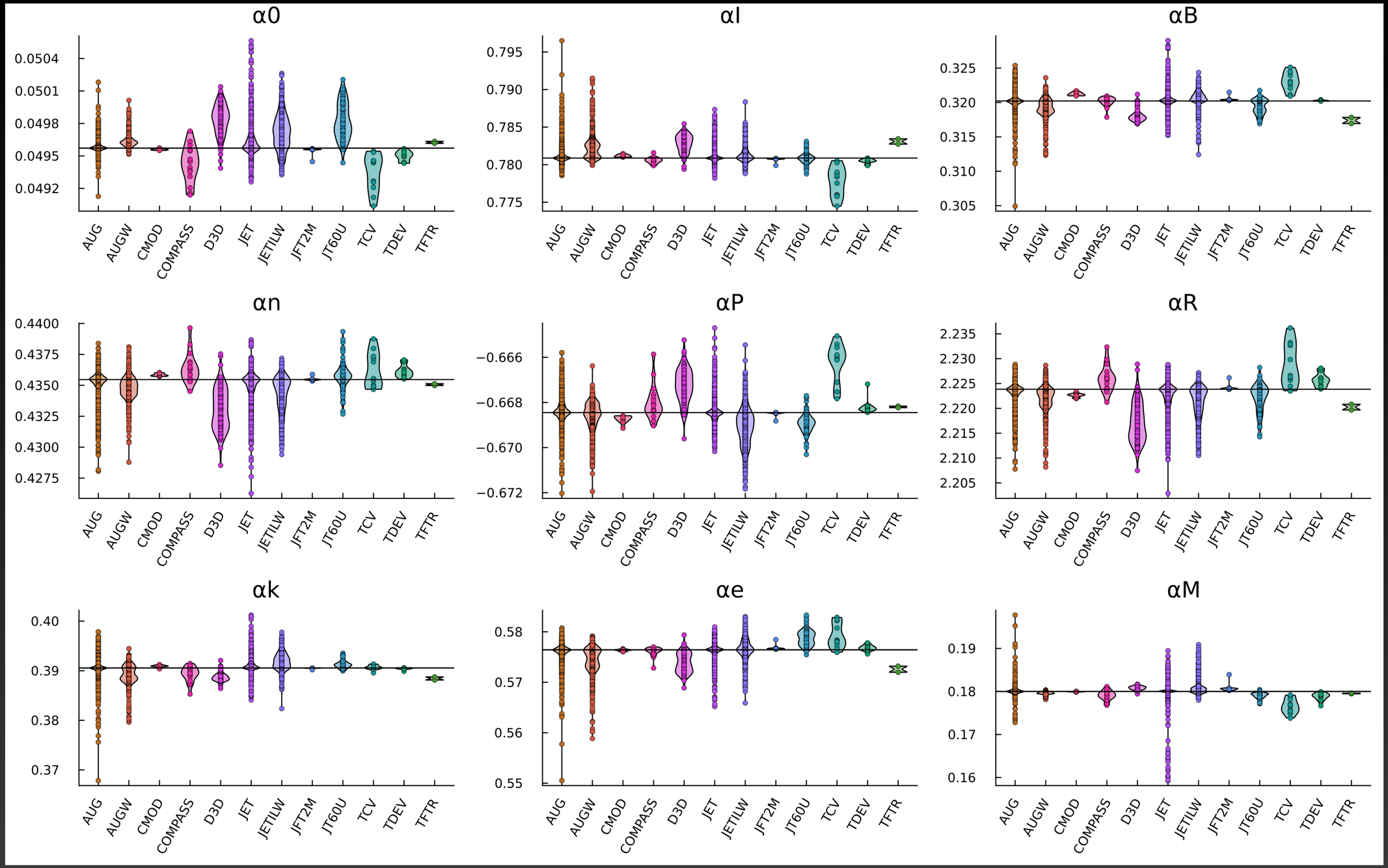


αR = 1.99

+

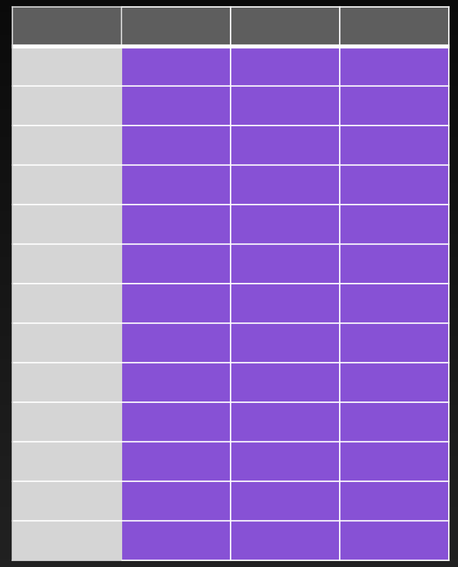
B/A




α0
0.0504
0.0501
0.0498
0.0495
0.0492
αI
0.795
0.790
0.785
0.780
0.775
αB
0.325
0.320
0.315
0.310
0.305
αn
0.4400
0.4375
0.4350
0.4325
0.4300
0.4275
αP
−0.666
−0.668
−0.670
−0.672
αR
2.235
2.230
2.225
2.220
2.215
2.210
2.205
αk
0.40
0.39
0.38
0.37
αe
0.58
0.57
0.56
0.55
αM
0.19
0.18
0.17
0.16
AUG
AUGW
CMOD
COMPASS
D3D
JET
JETILW
JFT2M
JT60U
TCV
TDEV
TFTR

Optimisation problem

No deviation set

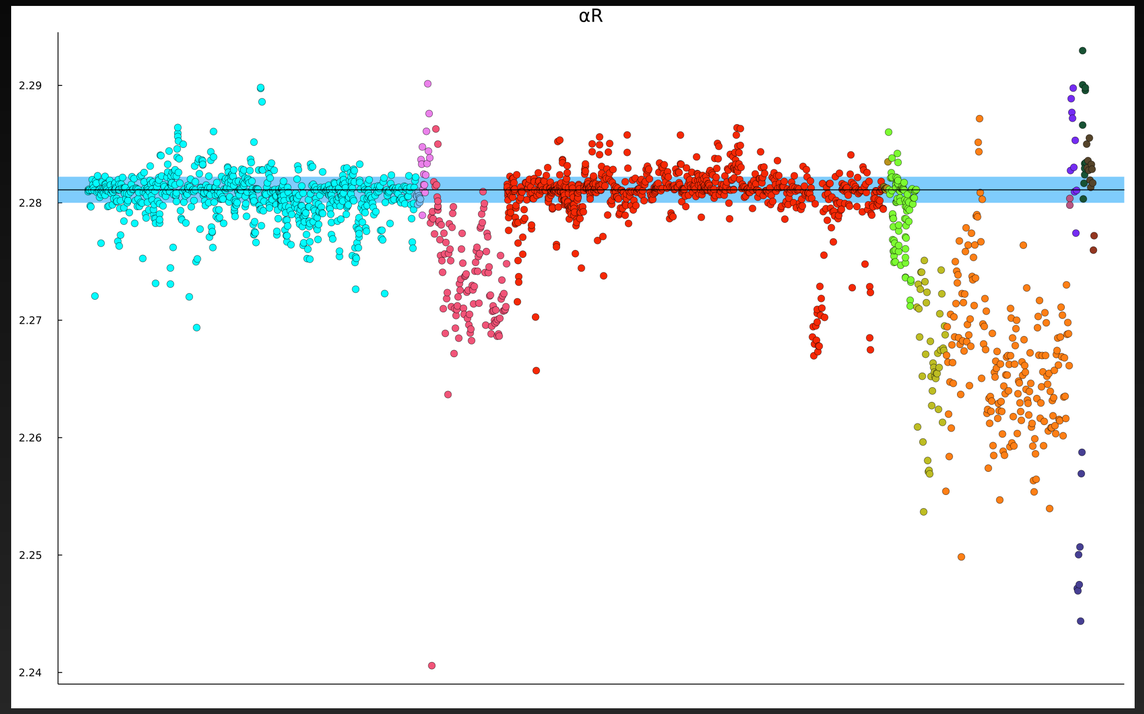


ℓ_1

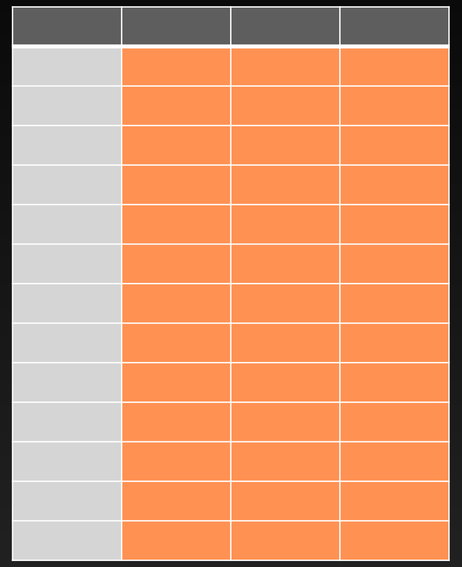
\+

A∩B

αR_{no_dev}



Deviation set



ℓ_2

\+

A∩B

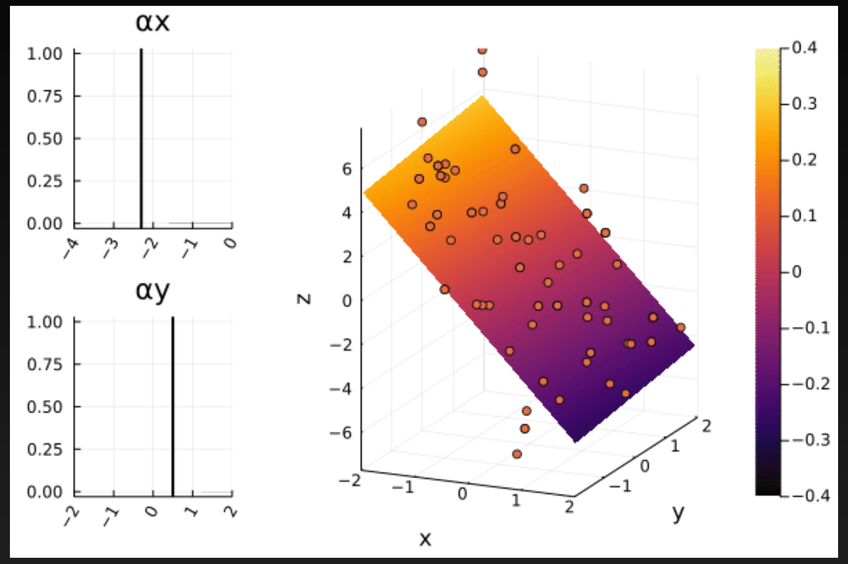
αR_{dev}

Objectives

Minimize ℓ_2

Maximize $|\alpha_{R_{no\_dev}} - \alpha_{R_{dev}}|$ s.t $\alpha_{R_{dev}} < \alpha_{R_{no\_dev}}$

Multi-collinearity



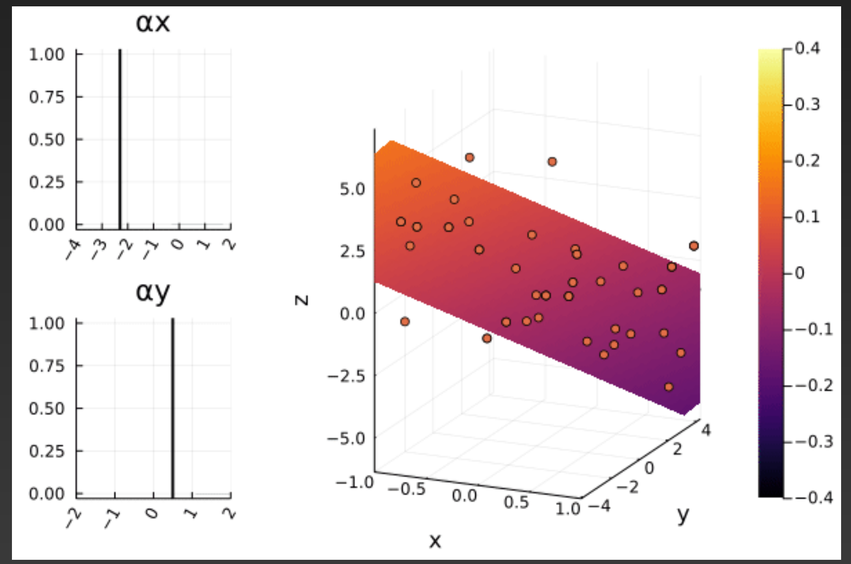
Uncorrelated clusters

- N (dimensionality of predictor space) clusters provide sufficient data to determine a pseudo stable regression hyperplane

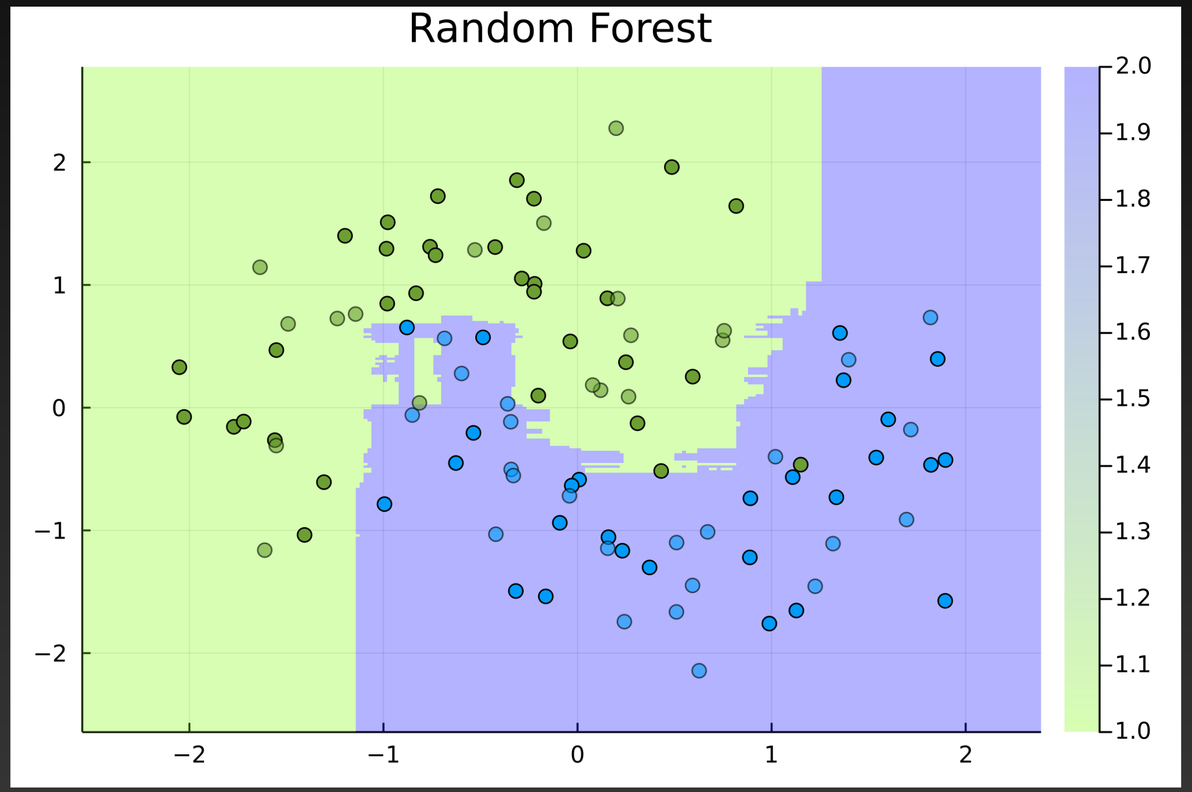
Correlated variables

$$y = 2x$$

- It is clear that the smallest deviations in data cause wild fluctuations in coefficients



Classification



- Can we find a well defined boundary in a (to be determined) feature space
- Can we make sense of that boundary
  - Physically
  - With a data driven relation