1)

a) Optimal Values:

l0_norm: 5

Minimum Validation Error: 2.3729 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x: 0.0063 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x_e: 0.0057 (Lambda: 5.0000, Index: 11)
Minimum Variance Error: 16.0708 (Lambda: 5.0000, Index: 11)

l0_norm: 10

Minimum Validation Error: 2.3232 (Lambda: 2.0000, Index: 10)
Minimum RMSE_x: 0.0073 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x_e: 0.0067 (Lambda: 5.0000, Index: 11)
Minimum Variance Error: 126.5253 (Lambda: 5.0000, Index: 11)

l0_norm: 15

Minimum Validation Error: 4.0873 (Lambda: 2.0000, Index: 10)
Minimum RMSE_x: 0.0065 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x_e: 0.0057 (Lambda: 5.0000, Index: 11)
Minimum Variance Error: 133.7211 (Lambda: 5.0000, Index: 11)

l0_norm: 20

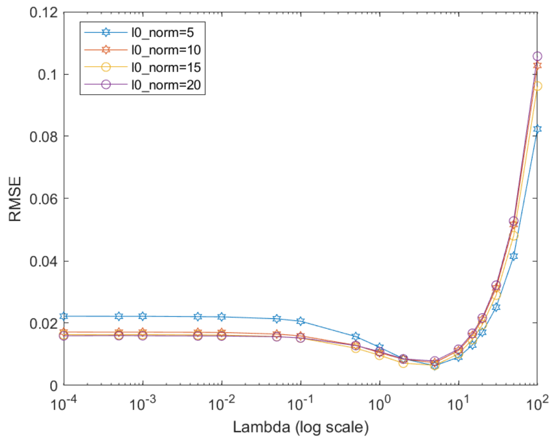
Minimum Validation Error: 4.7170 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x: 0.0079 (Lambda: 5.0000, Index: 11)
Minimum RMSE_x_e: 0.0074 (Lambda: 5.0000, Index: 11)
Minimum Variance Error: 178.1549 (Lambda: 5.0000, Index: 11)

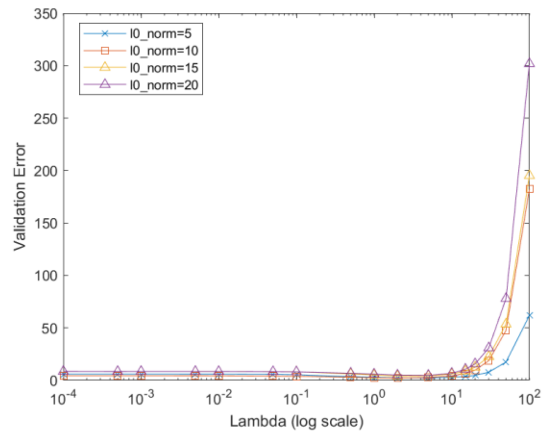
RMSE_x and RMSE_x_e are RMSE errors for case a), e)

The Optimal Values does not necessarily agree based on VE but the value is still close,

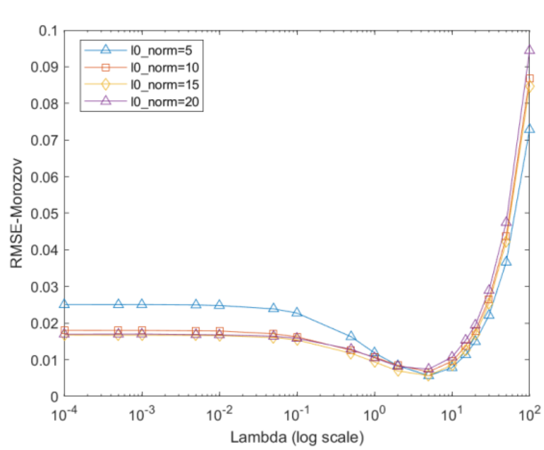
For Morozov’s method and RMSE they agree most of the time.

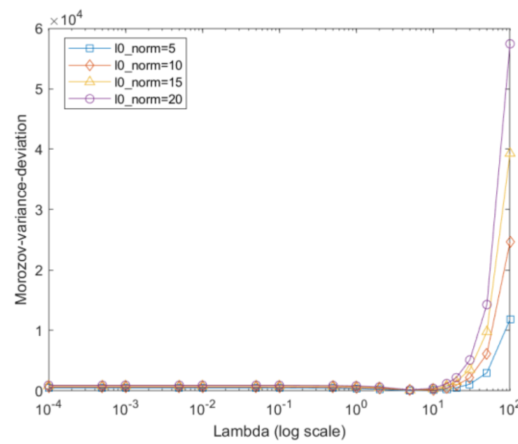
Cross Validation





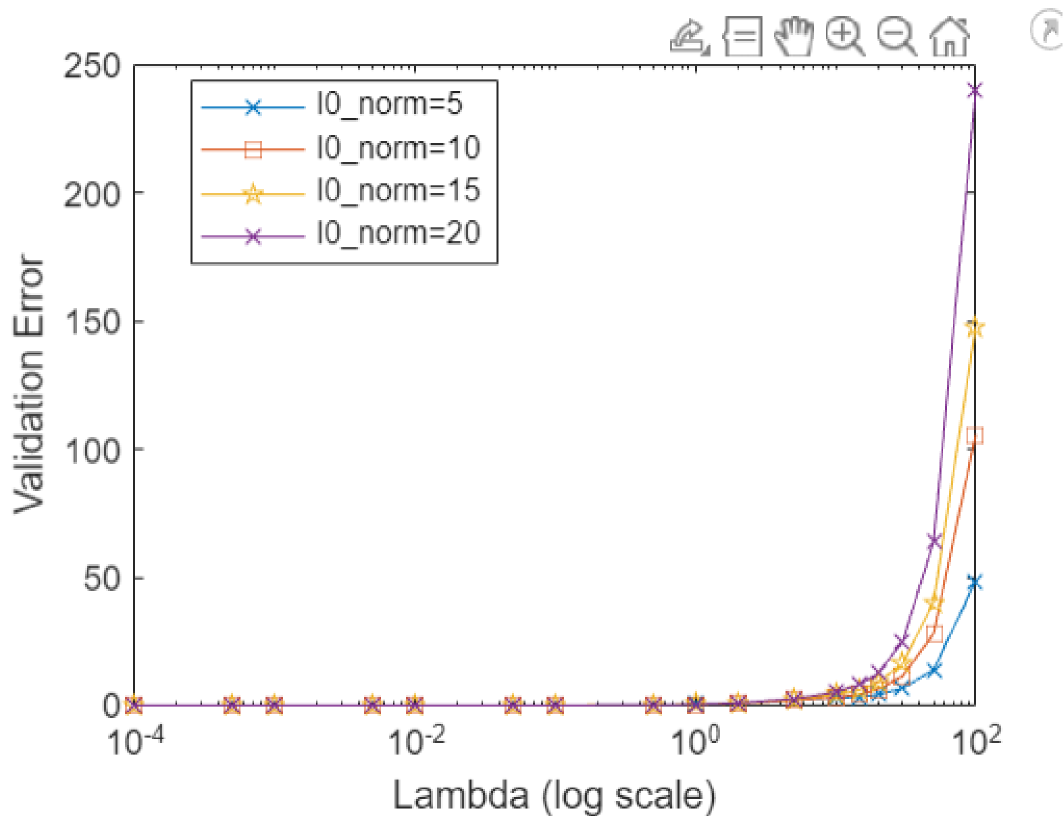
Morozov’s discrepancy principle





In all the plots, the y-axis values initially stay very close (mostly slightly decreasing with lambda) and then they increase with lambda

b)



Above is the graph for validation error if Reconstruction set and Validation set are same.

A random Lambda is selected since Validation error is zero for small values of lambda(<10).

Probably because the reconstructed vector was in the null space of the submatrix(of sensingMatrix*BasisMatrix) corresponding to the crossValidation Measurements.

For higher lambda we see Cross Validation error shoots up. Theoretical Guarantees don't exist in this case.

c)Theorem1

(photo taken from the paper https://ieeexplore.ieee.org/document/6854225)

Theorem 1. *(Recovery error estimation): Provided that m_{cv} is sufficiently large, with probability* $\mathrm{erf}(\frac{\lambda}{\sqrt{2}})$ *the following holds*

$$h(\lambda, +)\epsilon_{\mathrm{cv}} - \sigma_{\mathrm{n}}^2 \leq \varepsilon_{\mathrm{x}} \leq h(\lambda, -)\epsilon_{\mathrm{cv}} - \sigma_{\mathrm{n}}^2, \tag{3}$$

where $h(\lambda, \pm)$ is a function related to λ defined as

$$h(\lambda, \pm) \triangleq \frac{m}{m_{\mathrm{cv}}} \frac{1}{1 \pm \lambda\sqrt{\frac{2}{m_{\mathrm{cv}}}}}, \tag{4}$$

and $\mathrm{erf}(\mathrm{u})$ *is the error function of normal distribution,*

$$\mathrm{erf}(\mathrm{u}) \triangleq \frac{1}{\sqrt{\pi}} \int_{-\mathrm{u}}^{\mathrm{u}} \mathrm{e}^{-\mathrm{t}^2} \mathrm{dt}. \tag{5}$$

Theorem 1 provides an answer to Prob. 1 bounding the recovery error ε_{x} by the interval

$$\left[h(\lambda, +)\epsilon_{\mathrm{cv}} - \sigma_{\mathrm{n}}^2, h(\lambda, -)\epsilon_{\mathrm{cv}} - \sigma_{\mathrm{n}}^2\right]. \tag{6}$$

The difference of the upper bound and the lower bound,

$$\frac{m}{m_{\mathrm{cv}}} \frac{2\lambda\sqrt{2}}{\sqrt{m_{\mathrm{cv}}} - \frac{2\lambda^2}{\sqrt{m_{\mathrm{cv}}}}} \epsilon_{\mathrm{cv}}, \tag{7}$$

is roughly proportional to $1/m_{\mathrm{cv}}^{2/3}$ and becomes tighter as m_{cv} increases.

m_{cv}: Number of Cross-Validation Measurements
ϵ_{cv}: Cross Validation Error
ϵ_x : Real Mean Square Error
σ_n : Noise Standard Deviation

Knowing CV error we can bound Real MSE and the difference between this errors decrease as number of cross validation measurements increase.

d)As we can see from Theorem 1 in part c , we have both Upper bounds and Lower bounds in Cross Validation Approach whereas previously we could only get only an Upper bound on norm of difference of vectors as that theorem only provided an lower bound on lambda

$$\lambda_N \geq 2\frac{\|\boldsymbol{X}^T\boldsymbol{w}\|_\infty}{N}$$

e)

Check a) for figures

Disadvantage: Morozov's discrepancy principle (MDP) needs information about noise variance whereas Cross-Validaiton (CV) is purely data-driven.

Advantage: the reconstruction in MDP uses all measurements whereas reconstruction in CV is done using only reconstruction set which is smaller.

It also takes into account the noise information and can potentially perform better.