

Visual Based Tuning of Regularized Kalman Filter for System Identification Problem

M.D. Grebenkin¹

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)

¹ ORCID: 0000-0002-3644-5912, mdgrebenkin@mephi.ru

Abstract

System identification problem for linear and non-linear systems utilizes a large set of algorithms for estimating a vector of model parameters, relying on measurements and system dynamics. In particular one can use a family of Kalman filter adaptive algorithms. In situation when system of interest is ill-conditioned it is proper to use regularized modification of Kalman filter. In comparison to standard algorithm, properly tuned RKF is significantly more stable to ill-conditioned problems, which frequently arise in the field of system identification due to limited observability or controllability of systems of interest.

This paper shows an approach for preliminary tuning and analyzing regularized Kalman filter algorithm (RKF) for parameter identification of a vector meter unit using visualization of its crucial values on a computer model. Visual based approach to RKF tuning on a computer model allows for simple and intuitive way to find suboptimal regularization strength and set it at initialization stage avoiding the necessity to include computationally expensive methods of real-time tuning in algorithm loop. It is shown that regularization strength value, found using this approach, yielded a better estimation accuracy not only in comparison with standard Kalman filter but in comparison with other possible regularization strength values as well.

Keywords: regularized Kalman filter, ill-conditioned problem, Tikhonov regularization, system dynamics visualization.

1. Introduction

Kalman filter algorithm was introduced by Rudolf Kalman in 1960 and this original implementation is still widely used in many areas of engineering and automatic system design. It is a discrete linear adaptive filter, that utilizes bayesian estimation of state vector of the dynamic linear system. If noise present in measurements is unbiased Gaussian, it's estimated vector is also unbiased and has minimal variance [1].

Kalman filter allows to take into account stochastic parameters of measurement noise, system's intrinsic dynamics, reaction to control signal and to minimize estimation bias and variance. Original Kalman filter algorithm does not require to tune any parameters except initial mean and variance of state vector and covariance matrices for measurement and state noise. However with the broadening range of applications many modifications of KF were introduced, including regularized Kalman filter variant [2, 3, 4], which purpose is to overcome effects of ill-conditioned measurements. Difficulties with ill-conditioned measurements frequently arise in system parameters identification area and lead to biased estimations and filter instability. The source of ill-conditioning may be bad observation matrix and limited or bad choice of controlling input signals while taking measurements.

Paper [2] substantiates mathematical concept of regularized Kalman filter and introduces its implementation with real-time regularization parameter tuning. This allows to accurately and dynamically adapt algorithm to incoming measurements and changes in stochastic state matrices, however for some application this may prove to be computationally expensive. In

such cases it is more suitable to choose suboptimal constant value for regularization parameter. It can be done preliminary by visualization and analyzing algorithm performance on computer model. This paper introduces new visual based approach to such tuning for problem of parameter identification of abstract 3 component vector meter unit, which takes known positions in 3D-space and measures reference vector value in form of 3 vector components and it's norm.

Computer model and algorithms were implemented in Matlab. All visualization made with matplotlib package for Python.

2. Regularized Kalman filter

Standard discrete Kalman filter estimates a state vector of linear dynamic system

$$\begin{aligned} x_{k+1} &= \Phi_{k+1,k}x_k + B_k u_k + w_k \\ z_{k+1} &= H_{k+1}x_{k+1} + V_{k+1} \end{aligned} \quad (1)$$

Here for each timestep k x_{k+1} is system state vector; w_k is unbiased Gaussian noise with covariance matrix Q ; z_{k+1} - measurement vector; V_{k+1} - unbiased Gaussian measurement noise with covariance matrix R ; $\Phi_{k+1,k}$ state innovation matrix from k to $k+1$ step; H_{k+1} is observation matrix, u_{k+1} is controlling signal vector and B_k its transformation matrix.

However for model parameter identification problem this system can be simplified. Parameters of vector meter unit are components of x_{k+1} and assumed to be constant w.r.t time, thus $\Phi_{k+1,k} = I$ and $w_k = 0$. Observation matrix H_{k+1} defined by rotation control program and known at each step k . Controlling signal does not affect vector of parameter, thus $u_{k+1} = 0$.

Kalman filter approximates probability distribution of x_{k+1} by calculating its conditional mean value $\hat{x}_{k+1|k+1}$ and conditional covariance matrix $P_{k+1|k+1}$ given sequence of measurements z_1, z_2, \dots, z_{k+1} obtained through $k+1$ steps. To do this algorithm carries out following calculations on each step:

1. Estimates a priori mean $\hat{x}_{k+1|k}$ and covariance $P_{k+1|k}$ state given measurements from previous steps

$$\begin{aligned} \hat{x}_{k+1|k} &= \Phi_{k+1,k} \cdot \hat{x}_{k|k} \\ P_{k+1|k} &= \Phi_{k+1,k} P_{k|k} \Phi_{k+1,k}^T + Q \end{aligned} \quad (2)$$

This step is omitted in particular problem, shown in this paper, since $\Phi_{k+1,k} = I$ and $w_k = 0$.

2. Calculates measurement residuals on current step given newly obtained measurements z_{k+1} :

$$dz_{k+1} = z_{k+1} - H_{k+1} \hat{x}_{k+1|k}. \quad (3)$$

3. Calculates Kalman gain matrix K_{k+1} :

$$\begin{aligned} K_{k+1} &= P_{k+1|k} H_{k+1}^T S^{-1} \\ S_k &= [H_{k+1} P_{k+1|k} H_{k+1}^T + R] \end{aligned} \quad (4)$$

4. Updates posterior mean $\hat{x}_{k+1|k+1}$ and covariance $P_{k+1|k+1}$ with measurement residuals and Kalman gain on step $k+1$:

$$\begin{aligned} \hat{x}_{k+1} &= \hat{x}_{k+1|k} + K_{k+1} dz_{k+1} \\ P_{k+1} &= (I - K_{k+1} H_{k+1}) P_{k+1|k} \end{aligned} \quad (5)$$

Algorithm starts with initial estimated mean \hat{x}_0 , covariance P_0 of state vector and measurement noise covariance R . Kalman filter gradually transforms posterior state covariance matrix $P_{k+1|k+1}$ to diagonal form and minimizes its trace [3].

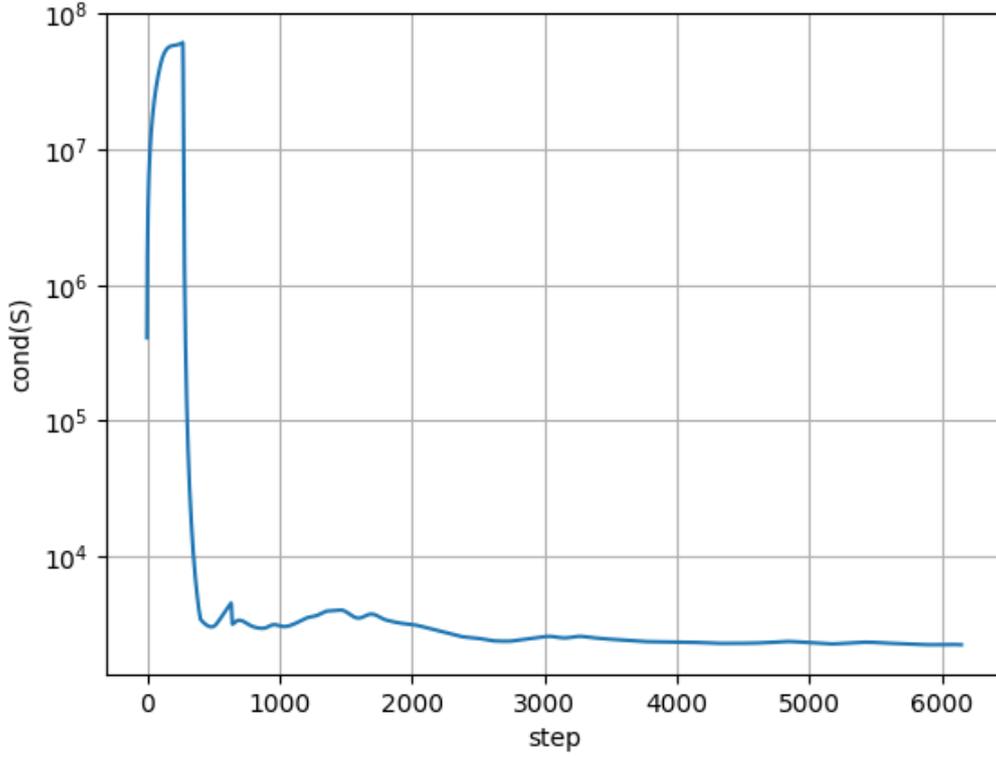


Figure 1. Condition number for matrix S_k .

Figure 1 shows, that despite condition number of $S_k = [H_{k+1}P_{k+1|k}H_{k+1}^T + R]$ quickly stabilizes, it is still too large for S_k to be considered good-conditioned.

Ill-conditioning affects stability of term $[H_{k+1}P_{k+1|k}H_{k+1}^T + R]^{-1}$ in (4) which leads to divergent estimates when measurement noise conditions deviate from unbiased Gaussian.

This issue is addressed in [2] by showing that Kalman gain can be as well calculated through inverting normal matrix [2, 5]

$$N_k = \Phi_{k,k+1}^T R^{-1} \Phi_{k,k+1} + P_{k|k+1}^{-1} \quad (6)$$

$$K_{k+1}^R = [\Phi_{k,k+1}^T R^{-1} \Phi_{k,k+1} + P_{k|k+1}^{-1}]^{-1} \Phi_{k,k+1}^T R^{-1}$$

which in its turn may be regularized via Tikhonov regularization to improve normal matrix conditioning. Tikhonov regularization of normal matrix is done by adding new term $\alpha_R A_k$ into part, that is inverted in (6):

$$K_{k+1}^R = [\Phi_{k,k+1}^T R^{-1} \Phi_{k,k+1} + P_{k|k+1}^{-1} + \alpha_R A_k]^{-1} \Phi_{k,k+1}^T R^{-1} \quad (7)$$

Here α_R is a scalar value known as regularization parameter and A_R is regularization matrix which should include priory filter information. For this reason $P_{k+1|k}^{-1}$ can be used as computationally cheap way to dynamically set A_R along all algorithm steps [2]. In order to keep computational complexity low regularization parameter α_R can be set on initialization stage and remain constant throughout all steps.

3. Choosing regularization parameter by performance visualization

In this work calibration parameter α_R is chosen with computer modeled visualization of Kalman filter performance. Parameter vector (state vector) x has 9 components and at each step 4 measurements are obtained, thus $dim(z_k) = 4 \times 1$. Estimation sequence has 6150 steps with every $H_k, k = 1..6150$ defined by control vectors for measuring unit.

Regularization parameter value is expected to vary in wide range for different implementation of (1) and sequences of H_k . It is proper to first visualize RKF performance for set of multiple α_R , increasing exponentially in power of 10 starting with 10^0 .

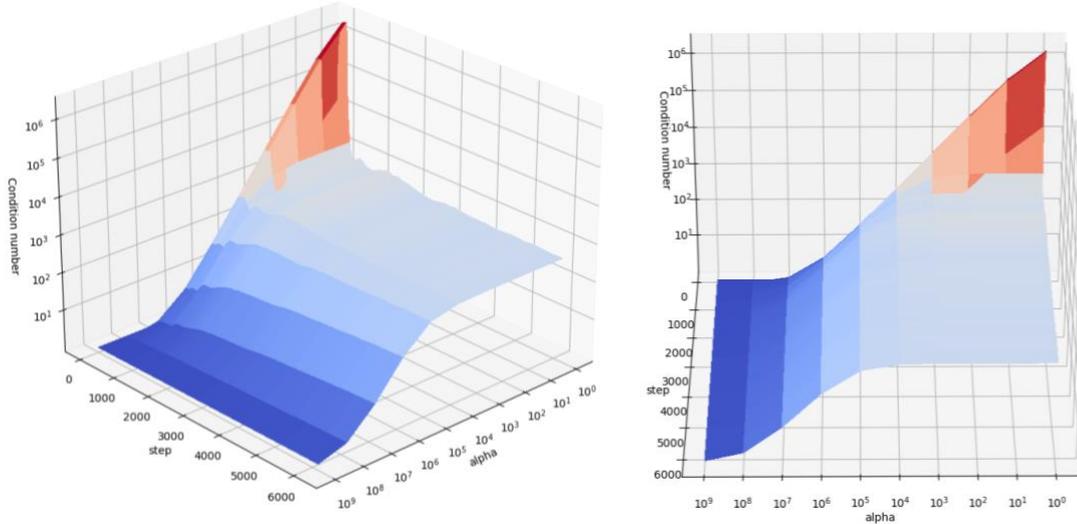


Figure 2. Change of conditioning for exponential set of α_R along all steps.

Figure 2 shows, that regularization starts to take effect only with $\alpha_R \approx 10^4$. In overall, conditioning monotonously depends on α_R and at values of regularization parameter around 10^7 it reaches values lesser that 100 by the end of the sequence.

However low condition number does not inherently mean better parameter estimation in terms of lesser error norm $\|\hat{x}_{end} - x\|$ and this is shown further.

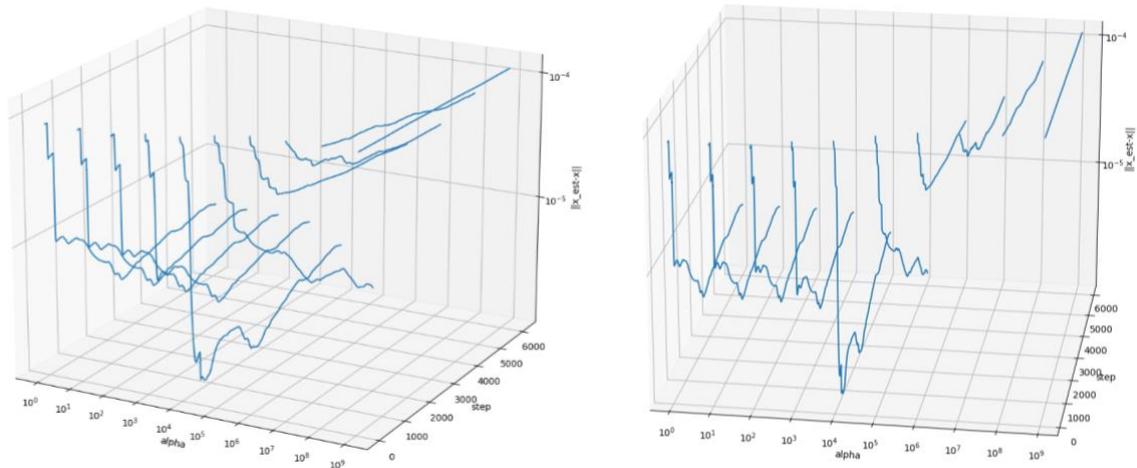


Figure 3. Change of estimation error norm for exponential set of α_R along all steps.

Figure 3 shows that despite best conditioning was archived with largest parameter α_R in set, the best estimation w.r.t norm $\|\hat{x}_{end} - x\|$ is actually obtained with much lesser $\alpha_R \approx 10^5$. The reason for this is that regularization term $\alpha_R A_k$ grows too large and effectively overpowers any other terms in N_k to the point when actual filter information form matrices H_k and R is not recoverable. Thus, the largest values of α_R correspond to biggest errors in actual estimation.

With optimal α_R localized in range $(10^4; 10^6)$ it is possible to increase precision in determining proper α_R . This time we use set of α_R with linear growth from 10^4 to 10^6 .

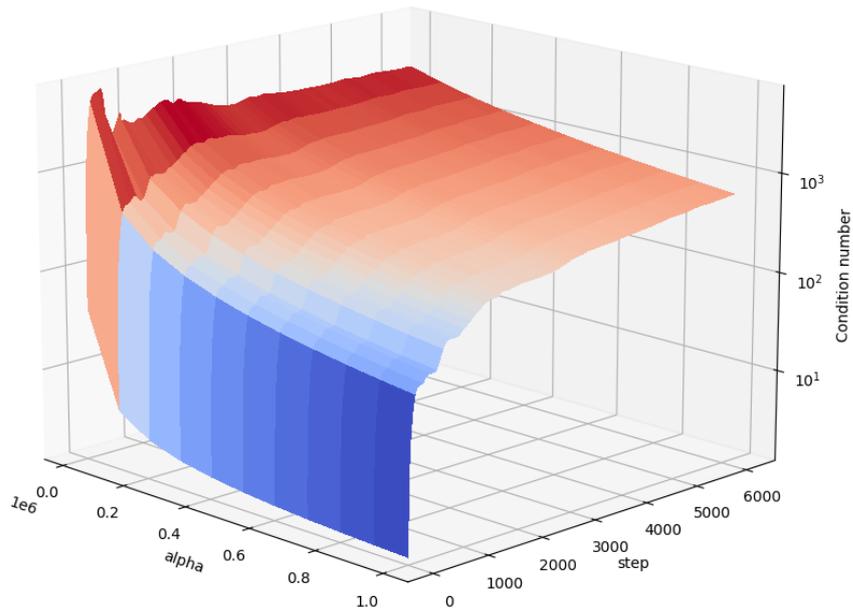


Figure 4. Change of conditioning for linear second set of α_R along all steps.

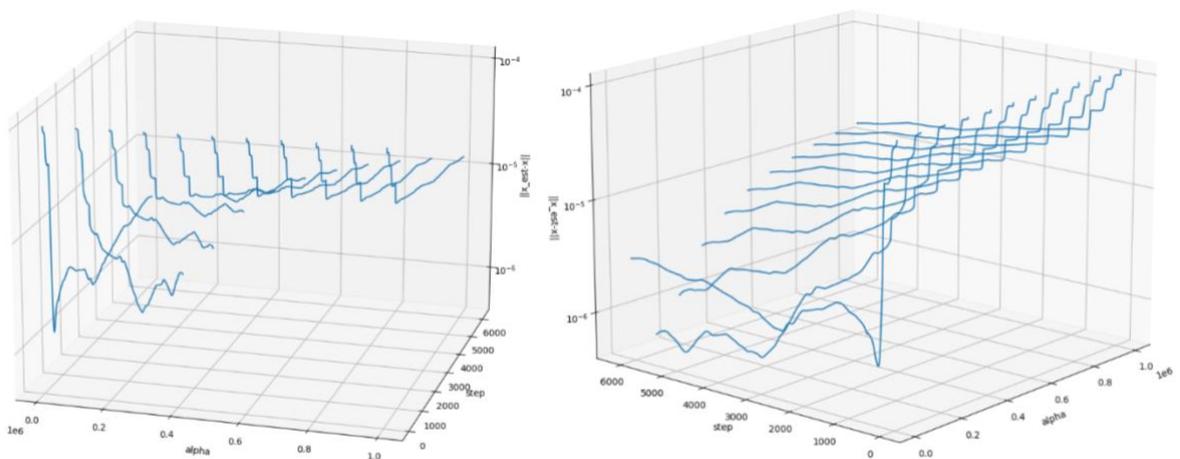


Figure 5. Change of estimation error norm for linear second set of α_R along all steps.

Figures 4 and 5 show that for α_R in range $(10^4; 10^6)$ condition number and estimation error norm have exponential dependence of α_R but behave in opposite to each other, which demonstrates an effect of excessively large regularization term. Best estimation in this set is archived with $\alpha_R = 109000$ which is still close to the result, obtained earlier with exponential set of regularization parameters.

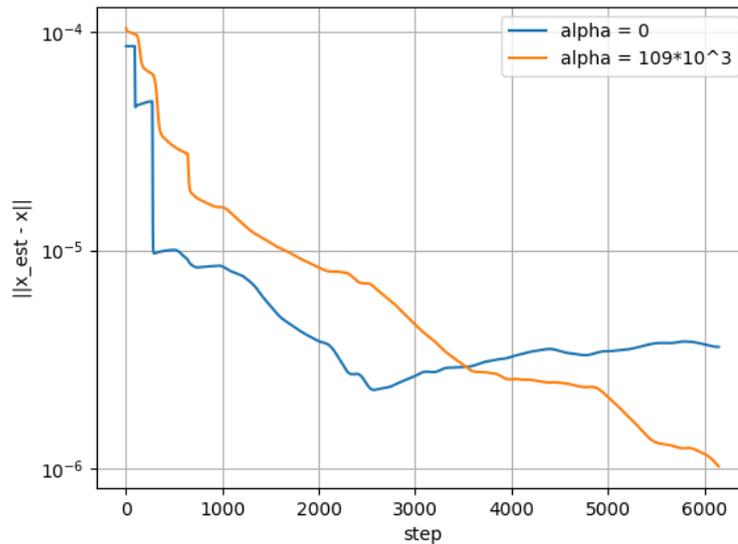


Figure 6. Change of estimation error norm for cases without regularization ($\alpha_R = 0$) and with regularization ($\alpha_R = 109 \cdot 10^3$) along all steps.

Table 1. Vectors of estimated parameters after full sequence. With and without regularization.

α_R	Relative error of estimation after full sequence, %									$\ \hat{x}_{end} - x\ $
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	
0	19.4154	39.5286	0.1814	17.7025	14.2140	0.3282	14.7019	39.9001	0.2408	$36.0771 \cdot 10^{-7}$
$109 \cdot 10^3$	15.2200	8.1408	0.1669	13.5619	6.1334	0.3179	5.6683	8.5110	0.2261	$5.2848 \cdot 10^{-7}$

Figure 6 and table 1 demonstrate effect of applying regularization to Kalman filter estimation problem. Although vector estimation made without regularization has faster convergence, ill-conditioning eventually makes estimation diverge. Vector estimation done with applied regularization term ($\alpha_R = 109 \cdot 10^3$), determined through visual analysis has slower but stable convergence. This yields to a smaller relative estimation error, overall norm of estimation error is almost 7 times less for vector estimated with RKF. Plot shows that by the end of sequence, estimation vector has not stopped converging. Longer measurement sequence might result in even smaller estimation error.

4. Conclusion

Kalman filter modifications may be used in conditions where application of original version of algorithm is not possible or too inefficient. Regularized modification of Kalman filter is purposed to avoid difficulties that arise with ill-conditioned problems. However it requires tuning in order to be effective.

Paper introduces a new approach for tuning regularized modification of linear discrete Kalman filter adaptive algorithm for system identification problem on a computer model via visualization of its performance and visual analysis of results This approach allowed to determine suboptimal constant regularization parameter and improve estimation precision not only in comparison with standard version of filter but also with other values of regularization parameter in range $(1; 10^9)$, since excessively large values render algorithm completely incoherent with a system, which parameters are identified. Suboptimal tuning yielded in smaller relative errors for components of estimated vector separately and in almost 7 times lower overall estimation error norm in comparison with standard filter version.

References

1. Kalman R. E. A New Approach to Linear Filtering and Prediction Problems, Transaction of the ASME-Journal of Basic Engineering, Vol.82, No.1, 1960, pp. 35-45.
2. Li Y., Gui Q., Han S., Gu Y.W. Tikhonov Regularized Kalman Filter and its Applications in Autonomous Orbit Determination of BDS. WSEAS Transactions on Mathematics, Vol. 16, 2017, pp. 187-196.
3. Milde T., Leistriz L., Astolfi L., Miltner W. H.R., Weiss T., Babiloni F., Witte H. A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials, NeuroImage, Volume 50, Issue 3, 2010, pp. 960-969
4. Golub G. H., Von Matt U. Tikhonov regularization for large scale problems. – Scientific Computing and Computational Mathematics Program, Computer Science Department, Stanford University, 1997, pp. 3-26.
5. Masnadi-Shirazi H., Masnadi-Shirazi A., Dastgheib M. A., A Step by Step Mathematical Derivation and Tutorial on Kalman Filters. arXiv e-prints, 2019, arXiv: 1910.03558.
6. J. Kaipio and E. Somersalo, Nonstationary Inverse Problems and State Estimation, Journal of Inverse and Ill-posed Problems, Vol7, No.3, 1999, pp. 273-282.
7. T. Lalith Kumar, K. Soundararajan. Performance comparison of adaptive filters with Kalman filter for speech enhancement, Bookman International Journal of Electrical & Electronics Engineering, N^o01, 2012, pp. 5-10.
8. Jigan V.I. Mnogoobrazie algoritmov adaptivnoy filtracii po kriteriyu naimenchikh kvadratov, Sovremennaya elektronika, N^o03, 2008, pp. 32-39.