FACTORIZED KALMAN FILTERING

Evgenia Suzdaleva

Department of Adaptive Systems
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Prague, CZECH REPUBLIC
E-mail: `suzdalev@utia.cas.cz`

Abstract: The paper deals with the Kalman filtering in the factorized form. The target application area is the urban traffic control, which main controlled variable – queue length, expressing the optimality of a traffic network most adequately, can not be directly observed and has to be estimated. Additional problem is that some state variables are of a discrete-valued nature. Thus, estimation of mixed-type data (continuous and discrete valued) models is highly desirable. A potential solution to this problem calls for a factorized version of the state-space model, which describes respective state factors individually. The present work considers the problem of the factorized filtering with Gaussian models and offers the solution, based on applying the $L'DL$ decomposition of the covariance matrix. The result of such a filtering is the posterior state estimate with the mean value and the factorized matrix of covariance.

Keywords: State Estimation, Factorized Filters, Traffic Control

1. INTRODUCTION

The paper deals with the Kalman filtering in the factorized form. The research in the area of factorization of the Kalman filter generated many results, which are worthy of notice. Only to enumerate some of them, one can note, for example, the following works. The paper (Dimitriu, 2005) describes the factorization of the covariance matrix in Kalman filter, where the covariance matrix was decomposed with the help of square root factorization. The QR-factorized filter and smoother algorithms for use on linear time-varying discrete-time problems, that can handle the general case of a singular state transition matrix, are discussed in (Psiaki, 1999). The UD-factorization of Kalman filter for the multi-sensor data fusion is presented in (Girija *et al.*, 2000). Another work, devoted to the UD-factorized covariance filter application, is concerned with development of a connected element interferometer (Morrison *et al.*, 2002). The method for particle filtering, which factorizes the likelihood, was proposed in (Patras and Pantic, 2004). It considers the problem, when the state space can be partitioned in groups of random variables, whose likelihood can be independently evaluated. As regards the nonlinear estimation, the following research works should be noted here. The square root form of unscented Kalman filter (UKF) for the state and parameter estimation, which, in its turn, was proposed as an alternative to the extended Kalman filter, used for nonlinear estimation, is described in (van der Merwe and Wan, 2001). This square-root UKF has better numerical properties and guarantees positive semi-definiteness of the underlying state covariance. The factorization of the covariance matrices is also used in problems of systems classification, dealing with multivariate Gaussian random field (Saltyte-Benth and Ducinskas, 2003).

Inspite of the variety of the mentioned works, their global aim was, primarily, to increase the numerical stability of the Kalman filter. The objective of the present work is *to obtain the estimates of the individual state entries*. The solution to this problem in general is concerned with a factorized version of the state-space model, which would enable to model the factors of the state individually. Use of the Gaussian state-space model and Gaussian observations calls for the Kalman filter.

The target application area for the present research is the urban traffic control. The motivation for the work is absolutely clear: nobody is surprised by the congestions in the crossroads of the cities, when the modern powerful cars have to move slowly and inefficiently within permanently extending peak hours and waste the time. Extension of the traffic network is expensive and often impossible, especially in historical cities. To solve the problem, all available means can be exploited: starting from economical pressure, various regulative measures up to the modern, ideally adaptive, feedback control. One of the main controlled variables in traffic systems is a queue length, which expresses the optimality of a traffic network most adequately. It can not be directly observed and has to be estimated. At the same time, other state variables might be of a discrete-valued nature. In this way, *estimation of mixed-type data (continuous and discrete valued) models* is highly desirable.

The layout of the paper includes the following sections. Section 2 reminds the basic facts about the Kalman filter and provides necessary notations, used throughout the text. Section 3 describes the idea of the presented version of the factorized Kalman filtering along with the model and algorithm used. Section 4 demonstrates examples of application of the algorithm to the system with different dimensions of the state. The remarks in Section 5 close the paper.

2. BASIC FACTS OF KALMAN FILTER

Throughout the text the following notations are used:

x_t is a quantity x at the discrete time instant labelled by $t \in t^* \equiv \{1, \dots, \mathring{t}\}$;

$\mathring{x}$ denotes the number of members in the countable set x^* or the number of entries in the vector x;

x_t is directly unobservable state of the system;.

y_t is a measured output of the system;

u_t is an optional input of the system.

2.1 Model

Let's assume, that the system is described by the state-space model

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + \omega_t, & (1) \\
y_t &= Cx_t + Du_t + e_t, & (2)
\end{aligned}
$$

where ω_t and e_t are Gaussian white noises with zero mean values and covariances Q and R correspondingly.

2.2 Algorithm

Kalman filter (Welch and Bishop, 1995) includes the following two steps of equations.

Time updating, which predicts the state estimate ahead in time

$$
\begin{aligned}
\hat{x}_{t+1}^- &= A\hat{x}_t + Bu_t, & (3) \\
P_{t+1}^- &= AP_tA' + Q. & (4)
\end{aligned}
$$

Data updating, which corrects the predicted estimate by the actual measurements

$$
\begin{aligned}
K_{t+1} &= P_{t+1}^- C'(CP_{t+1}^- C' + R)^{-1}, & (5) \\
\hat{x}_{t+1} &= \hat{x}_{t+1}^- + K_{t+1}(y_t - C\hat{x}_{t+1}^- - Du_t), & (6) \\
P_{t+1} &= (I - K_{t+1}C)P_{t+1}^-. & (7)
\end{aligned}
$$

The result of the algorithm application is the normal distribution of the state with mean value $\hat{x}_{t+1}$ and covariance matrix P_{t+1}. The initial values $\hat{x}_0$, P_0 are known.

3. FACTORIZED KALMAN FILTERING

The state-space model, supposed to be used for the factorized Kalman filtering, has practically the same form as (1-2), with the exception of covariance matrices of noises:

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + H'\omega_t, & (8) \\
y_t &= Cx_t + Du_t + F'e_t. & (9)
\end{aligned}
$$

Here ω_t and e_t are Gaussian white noises, for which $f(\omega) \sim \mathcal{N}(0, Q)$, $f(e) \sim \mathcal{N}(0, R)$. $H'QH$ and $F'RF$ are $L'DL$ decomposed matrices, where L is lower triangular matrix with unit diagonal and D is a diagonal matrix. The matrix A is of dimension $(\mathring{x} \times \mathring{x})$, $B - (\mathring{x} \times \mathring{u})$, $C - (\mathring{y} \times \mathring{x})$ and $D - (\mathring{y} \times \mathring{u})$.

The state estimate is assumed to be calculated with $L'DL$-factorized covariance matrix, i. e. as $\mathcal{N}(\hat{x}_{t+1}; P_{t+1})$, where $P_{t+1} = L'_{P(t+1)}D_{P(t+1)}L_{P(t+1)}$. The initial values $\hat{x}_0, P_0 = L'_0D_0L_0$ are known. Now let's consider in details the time updating and the data updating procedures for the factorized filter.

3.1 Time updating

Let's calculate the prior state estimate for the time $t + 1$. Here we multiply equation (3) by inverse matrix H'^{-1}.

$$
H'^{-1}\hat{x}_{t+1}^- = H'^{-1}A\hat{x}_t + H'^{-1}Bu_t. \tag{10}
$$

Denote $H'^{-1} = G$, we obtain

$$
G\hat{x}_{t+1}^- = G \underbrace{A\hat{x}_t}_{\tilde{\hat{x}}_t} + \underbrace{GB}_{\tilde{B}} u_t = G\tilde{\hat{x}}_t + \tilde{B}u_t, \tag{11}
$$

where new state $\tilde{x}_t$ is a known vector, G is the upper triangular matrix with unit diagonal, matrix $\tilde{B}$ is of dimension $(\mathring{x} \times \mathring{u})$. The structure of matrices in equations (11) can be illustrated in the following way (for the case $\mathring{x} = 3$, $\mathring{u} = \mathring{y} = 2$).

$$
\underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\hat{x}_{t+1}^-} = \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\hat{x}_t} + \underbrace{\begin{bmatrix} \cdots & \cdots \\ \cdots & \cdots \\ \cdots & \cdots \end{bmatrix}}_{\tilde{B}} \underbrace{\begin{bmatrix} \cdot \\ \cdot \end{bmatrix}}_{u_t} =
$$

$$
= \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\tilde{x}_t} + \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\tilde{B}u_t}
$$

Considering equation (11) in the form of individual factors, we can present the predicted estimate for the i-th state factor with dimension of the state vector, equal to $\mathring{x}$, as

$$
\hat{x}^-_{i(t+1)} + \sum_{k=i+1}^{\mathring{x}} g_{ik}\hat{x}^-_{k(t+1)} = \tilde{\hat{x}}_{i(t)} + \sum_{k=i+1}^{\mathring{x}} g_{ik}\tilde{\hat{x}}_{k(t)} + \sum_{j=1}^{\mathring{u}} \tilde{b}_{ij}u_{j(t)}, \tag{12}
$$

$$
\hat{x}^-_{i(t+1)} = \tilde{\hat{x}}_{i(t)} + \sum_{k=i+1}^{\mathring{x}} g_{ik}(\tilde{\hat{x}}_{k(t)} - \hat{x}^-_{k(t+1)}) + \sum_{j=1}^{\mathring{u}} \tilde{b}_{ij}u_{j(t)}. \tag{13}
$$

The calculation of covariance matrix is done as

$$
P^-_{t+1} = AP_tA' + H'QH = AL'_{Pt}D_{Pt}L_{Pt}A' + H'QH = L'^-_{P(t+1)}D^-_{P(t+1)}L^-_{P(t+1)}. \tag{14}
$$

The illustration of the structure of matrices is the following one.

$$
\underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{L'_{Pt}} \underbrace{\begin{bmatrix} \cdots & & \\ & \cdots & \\ & & \cdots \end{bmatrix}}_{D_{Pt}} \underbrace{\begin{bmatrix} \cdots & & \\ \cdots & \cdots & \\ \cdots & \cdots & \cdots \end{bmatrix}}_{L_{Pt}} \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}}_{A'}^{'} +
$$

$$
+ \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{H'} \underbrace{\begin{bmatrix} \cdots & & \\ & \cdots & \\ & & \cdots \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \cdots & & \\ \cdots & \cdots & \\ \cdots & \cdots & \cdots \end{bmatrix}}_{H} =
$$

$$
= \underbrace{\begin{bmatrix} \cdots & \cdots & \cdots \\ & \cdots & \cdots \\ & & \cdots \end{bmatrix}}_{L'^-_{P(t+1)}} \underbrace{\begin{bmatrix} \cdots & & \\ & \cdots & \\ & & \cdots \end{bmatrix}}_{D^-_{P(t+1)}} \underbrace{\begin{bmatrix} \cdots & & \\ \cdots & \cdots & \\ \cdots & \cdots & \cdots \end{bmatrix}}_{L^-_{P(t+1)}}.
$$

Note, that the result in (14) is obtained as the $L'DL$-decomposed matrix of covariance. The calculation exploits the algorithms from the toolbox Mixtools (Nedoma *et al.*, 2003) and does not contain numerically dangerous operations.

3.2 Data updating

Let's begin untraditionally from calculation of the covariance matrix. The advantages of such a calculation will be clear later. The straightforward procedure of the covariance matrix calculation is based on the matrix inversion lemma (Peterka, 1981).

Proposition 3.1 (Matrix Inversion Lemma)

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \tag{15}$$

Proof: See (Peterka, 1981). □

The matrix of covariance is calculated in the step of data updating as

$$P_{t+1} = (I - P_{t+1}^- C'(CP_{t+1}^- C' + F'RF)^{-1}C)P_{t+1}^-, \tag{16}$$

where

$$P_{t+1}^- C'(CP_{t+1}^- C' + F'RF)^{-1} = K_{t+1} \tag{17}$$

is the Kalman gain matrix. We will need it later, but let's now return to equation (16) in the form, as it is presented. After multiplying P_{t+1}^- from the right of the brackets we obtain

$$
\begin{aligned}
P_{t+1} &= P_{t+1}^- - P_{t+1}^- C'(CP_{t+1}^- C' + F'RF)^{-1}CP_{t+1}^-, & (18) \\
&= P_{t+1}^- - P_{t+1}^- C'(F'RF + CP_{t+1}^- C')^{-1}CP_{t+1}^-, & (19) \\
&= (P_{t+1}^{-\,-1} + C'(F'RF)^{-1}C)^{-1}, & (20) \\
&= L'_{P(t+1)}D_{P(t+1)}L_{P(t+1)}. & (21)
\end{aligned}
$$

Proof: The relation (20) is calculated straightforward, based on application of the matrix inversion lemma. □

The covariance matrix is obtained in $L'DL$-factorized form, which can be illustrated similarly, as it was shown in Section 3.1. For calculating the algorithms from toolbox Mixtools (Nedoma *et al.*, 2003) are used.

Now we can return to the Kalman gain matrix, calculated in (17). It results in the matrix K_{t+1} with $\mathring{y}$ columns and $\mathring{x}$ rows, where $\mathring{y}$ is dimension of the output vector. With its help we can calculate the posterior state estimate as it follows.

$$\hat{x}_{t+1} = \hat{x}_{t+1}^- + K_{t+1}\underbrace{(y_t - \underbrace{C\hat{x}_{t+1}^- - Du_t}_{\text{denote by } \hat{y}_t})}_{\text{or denote by vector } \Delta_t}, \tag{22}$$

$$\hat{x}_{t+1} = \hat{x}_{t+1}^- + K_{t+1}\Delta_t, \tag{23}$$

where Δ_t is a known vector with $\mathring{y}$ rows. The schematic representation of (23) looks like

$$\underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\hat{x}_{t+1}} = \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\hat{x}_{t+1}^-} + \underbrace{\begin{bmatrix} \cdots & \cdots \\ \cdots & \cdots \\ \cdots & \cdots \end{bmatrix}}_{K_{t+1}} \underbrace{\begin{bmatrix} \cdot \\ \cdot \end{bmatrix}}_{\Delta_t},$$

Thus, we can consider the equation (23) from the point of view of estimates for the individual factors of the state. We have

$$\hat{x}_{i(t+1)} = \hat{x}^-_{i(t+1)} + \sum_{j=1}^{\mathring{x}} k_{ij}\Delta_{j(t)}, \tag{24}$$

where

$$\Delta_{i(t)} = y_{i(t)} - \sum_{j=1}^{\mathring{y}} c_{ij}\hat{x}^-_{j(t+1)} - \sum_{j=1}^{\mathring{u}} d_{ij}u_{j(t)}. \tag{25}$$

Summarizing the section, the suggested factorized Kalman filtering involves the following algorithm.

Time updating

$$H'^{-1}\hat{x}^-_{t+1} = H'^{-1}A\hat{x}_t + H'^{-1}Bu_t, \tag{26}$$
$$P^-_{t+1} = AP_tA' + H'QH, \tag{27}$$

Data updating

$$K_{t+1} = P^-_{t+1}C'(CP^-_{t+1}C' + F'RF)^{-1}, \tag{28}$$
$$P_{t+1} = (P^-_{t+1}{}^{-1} + C'(F'RF)^{-1}C)^{-1}, \tag{29}$$
$$\hat{x}_{t+1} = \hat{x}^-_{t+1} + K_{t+1}(y_t - C\hat{x}^-_{t+1} - Du_t). \tag{30}$$

4. EXAMPLES

Three examples of the factorized state estimation are shown at this section. 200 data, matrices A, B, C and D and covariances have been simulated for the state-space model with rather small noise, taking into account the dimension of the system. The first example is a system with single input, single output and single state. The results – the simulated and the estimated states (left) and prediction of the output (right) – are plotted in Figure 1. The second example uses
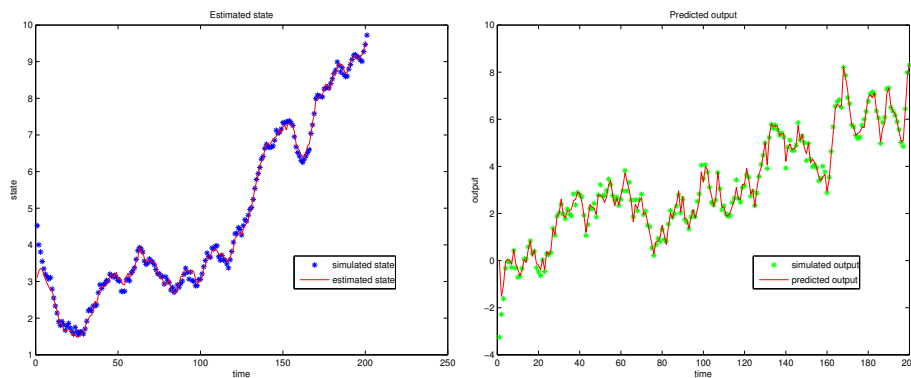

Fig. 1: Factorized Kalman filter for one-dimensional system

the system with two states, single input and single output. The estimation of two-dimensional states and predicted output can be seen at Figure 2. The last example provides the case with $\mathring{x} = 3$, $\mathring{u} = \mathring{y} = 2$, see Figure 3. The different initial values confirm the functioning of
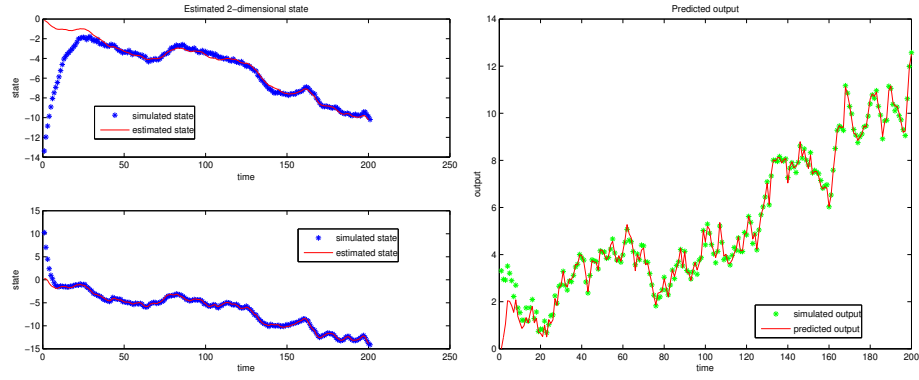
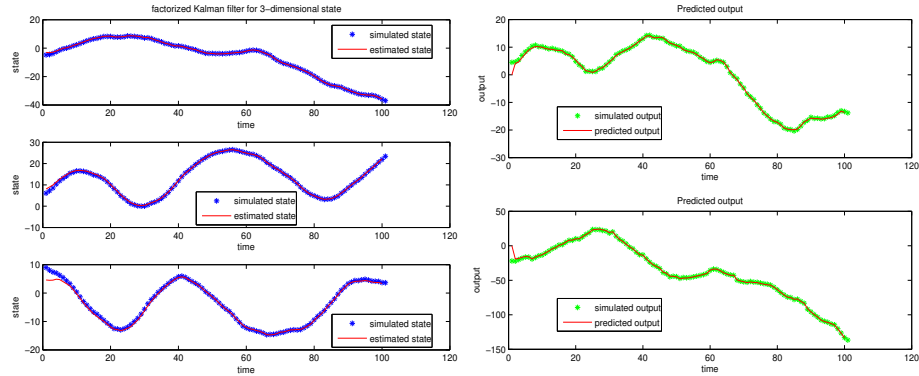Fig. 2: Estimation of two-dimensional state and output prediction

Fig. 3: Estimation of 3-dimensional state and prediction of 2-dimensional output

the filter. The results have been compared with the results, obtained from the Kalman filter. For all the cases the state mean value calculation gives the same quantity. The covariance matrices calculation (after multiplying the $L'DL$ decomposition) results in very close values (a difference is about $8.0085e - 17$). Results of the whiteness test (Wonnacott and Wonnacott, 1984) for the prediction error are given in Table 1. Increasing the noise, one can obtain the

Table 1: Whiteness test of the prediction error

Example	p-value
1	0.1975
2	0.0009071
3	0.0003164

higher values of the probability of the elements independence. Nevertheless, the results show, that there is no more information to be extracted from the sample.

5. CONCLUSIONS

The paper presents the factorized version of the Kalman filter. The described filtering is expected not only to contribute to higher numerical stability of the filter, but also to enable handling with the mixed-type (continuous and discrete valued) data. The present work demonstrates the specialization of the general solution of the factorized state estimation to linear

Gaussian model. The experiments with the mixed-type data will be the part of future work. Testing on realistic simulation of traffic control problem is also planned.

REFERENCES

Dimitriu, G. (2005). Implementation issues related to data assimilation using Kalman filtering. In: *Proceedings for 3th World Conference on Computational Statistics & Data Analysis*. International Association for Statistical Computing. Limassol, Cyprus.

Girija, G., J. R. Raol, R. Appavu Raj and S. Kashyap (2000). Tracking filter and multi-sensor data fusion. In: *Sadhana. Special Issue on Advances in Modelling, System Identification & Parameter Estimation*. Vol. 25. pp. 159–167. Indian Academy of Sciences.

Morrison, D., S. Pogorelc, T. Celano and A. Gifford (2002). Ephemeris determination using a connected element interferometer. In: *34th Annual Precise Time and Time Interval (PTTI) Meeting*. Reston, Virginia, USA.

Nedoma, P., M. Kárný, T.V. Guy, I. Nagy and J. Böhm (2003). *Mixtools (Program)*. ÚTIA AV ČR. Prague.

Patras, I. and M. Pantic (2004). Particle filtering with factorized likelihoods for tracking facial features. In: *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)*. Seoul, Korea.

Peterka, V. (1981). Bayesian approach to system identification. In: *Trends and Progress in System Identification (P. Eykhoff ed.)*. pp. 239–304. Pergamon Press. Oxford.

Psiaki, M. L. (1999). Square-root information filtering and fixed-interval smoothing with singularities. *Automatica* **35**(7), 1323–1331.

Saltyte-Benth, J. and K. Ducinskas (2003). Linear discriminant analysis of multivariate spatial-temporal regressions. Technical Report 2. Department of Mathematics, University of Oslo. ISBN 82-553-1395-8.

van der Merwe, R. and E. A. Wan (2001). The square-root unscented Kalman filter for state and parameter-estimation. In: *International Conference on Acoustics, Speech, and Signal Processing*. Salt Lake City, Utah.

Welch, Greg and Gary Bishop (1995). An Introduction to the Kalman Filter. Technical Report 95-041. UNC-CH Computer Science.

Wonnacott, Thomas H. and Ronald J. Wonnacott (1984). *Introductory statistics for business and economics*. 4th ed.. John Wiley. New York.