A Case Study on Traffic Matrix Estimation Under Gaussian Distribution

Ilmari Juva Pirkko Kuusela Jorma Virtamo

March 15, 2004

Abstract

We report a case study on an iterative method of traffic matrix estimation under some simplifying assumptions about the distribution of the origin-destination traffic demands. The starting point of our work is the Vaton-Gravey iterative Bayesian method, but there are quite a few differences between that method and our consideration. It is assumed that the distribution of the demands follow a single Gaussian distribution instead of being a modulated process. The normality assumption allows us to bypass the Markov Chain Monte Carlo step in the iterative method and explicitly derive the expected values for mean and covariance matrix of the traffic demands conditioned on link counts. We show that under the assumption of single underlying distribution the expected values of the mean and covariance converges after the first step of the iteration. This method cannot improve on this if no relation between mean and variance is imposed in order to make use of the covariance matrix estimates, or the distribution is assumed to be modulated from a regime of distributions.

1 Introduction

In many traffic engineering applications, the knowledge on the underlying traffic volumes is assumed. The *Traffic Matrix* gives the amount of demanded traffic between each node in the network. The traffic matrix can not be directly measured, so there are not yet many methods to obtain them, although it is recognized that accurate traffic demand matrices are crucial for traffic engineering.

The only information readily available are the link loads and the routing matrix. The traffic demands $\boldsymbol{x}$ between origin-destination pairs and the routing matrix $\boldsymbol{A}$ determine the link loads

$\boldsymbol{y}$ through relation

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{1}$$

Since in any realistic network there are more OD pairs than links, the problem of solving $\boldsymbol{x}$ from $\boldsymbol{A}$ and $\boldsymbol{y}$ is strongly underdetermined and thus explicit solutions cannot be found.

Most promising proposed methods for inferring traffic matrix from link loads include Bayesian based inference techniques and network tomography. Bayesian methods compute conditional probability distributions for elements of the traffic matrix, given the observed link loads. This method usually employs Markov-chain Monte Carlo simulation for computing the posterior probability. Network tomography uses more classical statistical methods like expectation maximization algorithm for calculating the maximum likelihood estimate for the traffic matrix based on the link loads.

The Vaton-Gravey method [1] consists of iteration and exchange of information between two boxes. The available data is the link counts on several successive time periods. The first box simulates the traffic matrix (OD counts) from the link counts utilizing some prior information on the OD counts at each fixed time period . As an example, the traffic counts for each OD pair are assumed to constitue a Markov modulated process. Then the successive values for each OD pair are fed into the second box that updates the parameters of the Markov modulated process, which are then fed back into the first box as a Bayesian prior and the process is repeated. The first box involves running a Markov Chain Monte Carlo simulation and the second box computes maximum likelihood estimate using the Expectation Maximization method.

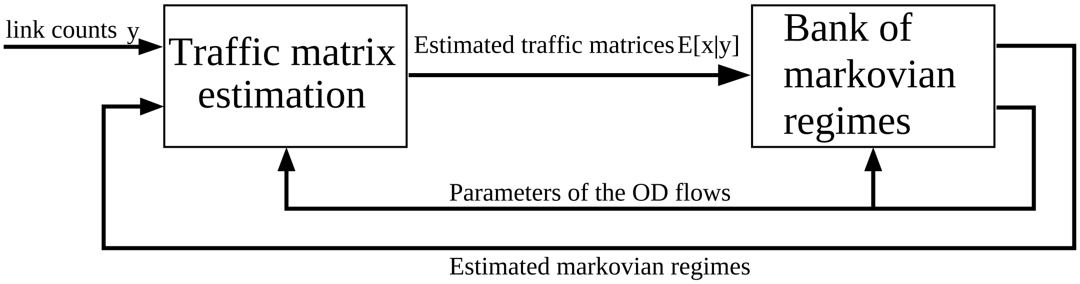

Figure 1: The Vaton-Gravey iterative method

In this report we consider explicitely a special case of the above idea. Our aim is to gain insight into the method and, in particular, into the output of the first box by examining a model that is simple enough to be computed analytically. We assume the OD pairs are independent and follow a single Gaussian distribution instead of a mixture of distributions. This reduces the complexity of the Vaton-Gravey method, and allows the explicit analysis. Our prior consists of the mean and the covariance matrix of the Gaussian distribution. The attractive feature of this approach is that the distribution conditioned on the link counts is again Gaussian. Thus we skip the MCMC simulation by calculating analytically the expected output of the first box.

Also we point out that, contrary to Vaton-Gravey method, in our approach the output of the first box is the whole conditioned distribution, not just the mean of an OD pair conditioned on the link count observation. This is because the means result in a distribution that is flattened out, and thus has a singular covariance matrix. See Figures 3-4 for illustration of the conditional "cloud" in our approach.

As in our case the output of the first box is the conditional distribution of $\boldsymbol{x}$ conditioned on $\boldsymbol{y}$, the second box has only the function of taking the expectation of these over $\boldsymbol{y}$. This yields the new estimate for the distribution of $\boldsymbol{x}$ which is here considered constant in time. The estimate is then returned to the first box as a prior. It turns out that the expected value of the mean does not change in the iteration after the first conditioning on link counts has been made. This result is proven later in the report.

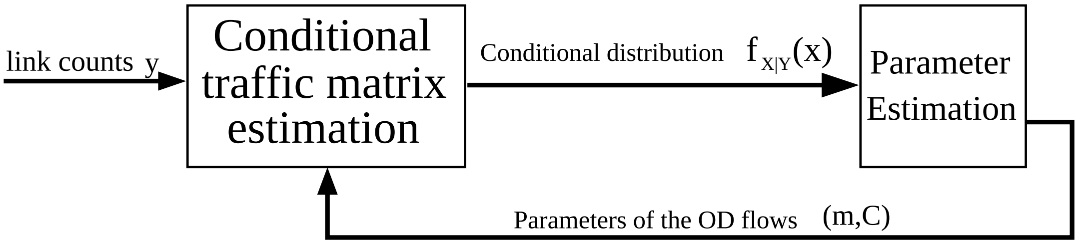

Figure 2: Illustration of the method studied.

The rest of the report is organized as follows: In section 2 the conditional distribution of $\boldsymbol{x}$ conditioned on $\boldsymbol{y}$ is derived. Then the the expected values for mean and covariance matrix estimates $(\boldsymbol{m}, \boldsymbol{C})$ are calculated. Section 3 illustrates the results of the previous section through example cases in a much simplified situation. In Section 4 we state that the iteration converges after the first step, and prove this result for the estimator of the mean. Section 5 gives some illustrative numerical examples, and in section 6 the report is summarized and some final conclusions made.

2 Conditional Gaussian distribution

2.1 Introduction

In this section we derive the equations for the conditional gaussian distribution. These are well known results but presented here for the sake of completeness. The distribution of variable $\boldsymbol{X}$, representing here the OD traffic amounts, is derived conditioned on $\boldsymbol{y}$, the measured link loads. From the conditional distribution we are able to solve the mean and the covariance matrix and their expected values. The covariance matrix is solved two ways. The conditional covariance approach is mathematically easier and more elegant, while the co-ordinate transformation method follows more the idea of the algorithm and is therefore probably more intuitively understandable.

2.2 The conditional distribution

Let the n-vector $\boldsymbol{x}$ represent a multivariate gaussian variable X with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{C}$,

$$f(\boldsymbol{x}) \sim \exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})). \tag{2}$$

Assume that we have a prior estimate $(\boldsymbol{m}, \boldsymbol{C})$ for $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We wish to determine the distribution of $\boldsymbol{X}$ conditioned on

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{y}, \tag{3}$$

where $\boldsymbol{y}$ in an m-vector and $\boldsymbol{A}$ is an $m \times n$ matrix with $m < n$. First we partition $\boldsymbol{x}$ and $\boldsymbol{A}$ as

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \qquad \boldsymbol{A} = (\boldsymbol{A}_1, \boldsymbol{A}_2),$$

where $\boldsymbol{x}_1$ is an m-vector, $\boldsymbol{x}_2$ is an $(n-m)$-vector, $\boldsymbol{A}_1$ in an $m \times m$-matrix, and $\boldsymbol{A}_2$ in an $m \times (n-m)$-matrix. From (3) we have

$$\boldsymbol{x}_1 = \boldsymbol{A}_1^{-1}(\boldsymbol{y} - \boldsymbol{A}_2\boldsymbol{x}_2). \tag{4}$$

By making the corresponding partition in the exponent of (2),

$$\begin{aligned} &(\boldsymbol{x}_1-\boldsymbol{m}_1)^{\mathrm{T}}\boldsymbol{B}_{11}(\boldsymbol{x}_1-\boldsymbol{m}_1) + (\boldsymbol{x}_1-\boldsymbol{m}_1)^{\mathrm{T}}\boldsymbol{B}_{12}(\boldsymbol{x}_2-\boldsymbol{m}_2) \\ &+(\boldsymbol{x}_2-\boldsymbol{m}_2)^{\mathrm{T}}\boldsymbol{B}_{21}(\boldsymbol{x}_1-\boldsymbol{m}_1) + (\boldsymbol{x}_2-\boldsymbol{m}_2)^{\mathrm{T}}\boldsymbol{B}_{22}(\boldsymbol{x}_2-\boldsymbol{m}_2), \end{aligned} \tag{5}$$

with

$$\boldsymbol{C}^{-1} = \boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix},$$

and substituting (4) into (5) we obtain

$$\begin{aligned}\boldsymbol{x}_2^{\mathrm{T}}\,(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{21}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2+\boldsymbol{B}_{22})\,\boldsymbol{x}_2 \\ +\boldsymbol{x}_2^{\mathrm{T}}((\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})(\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{m}_1)+(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\boldsymbol{m}_2) \\ +(\text{transpose})\boldsymbol{x}_2+\text{constant}.\end{aligned} \tag{6}$$

Because of symmetry the multiplier of x_2 is just the transpose of the multiplier of x_2^{T} We wish to write this as complete square of the form

$$(\boldsymbol{x}_2-\tilde{\boldsymbol{m}}_2)^{\mathrm{T}}\tilde{\boldsymbol{C}}_{22}^{-1}(\boldsymbol{x}_2-\tilde{\boldsymbol{m}}_2)+\text{constant}=\boldsymbol{x}_2^{\mathrm{T}}\tilde{\boldsymbol{C}}_{22}^{-1}\boldsymbol{x}_2-(\boldsymbol{x}_2^{\mathrm{T}}\tilde{\boldsymbol{C}}_{22}^{-1}\tilde{\boldsymbol{m}}_2+(\text{transpose})\boldsymbol{x}_2)+\text{constant}, \tag{7}$$

Now we can pick out the terms $\tilde{\boldsymbol{C}}_{22}^{-1}$ and $\tilde{\boldsymbol{C}}_{22}^{-1}\tilde{\boldsymbol{m}}_2$ from (6), which are the multipliers of the quadratic term and the $\boldsymbol{x}_2^{\mathrm{T}}$ term respectively.

$$\begin{aligned}\tilde{\boldsymbol{C}}_{22}^{-1} &= \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2-\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{(A_1^{-1})}^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{21}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2+\boldsymbol{B}_{22} \qquad (8) \\ \tilde{\boldsymbol{C}}_{22}^{-1}\tilde{\boldsymbol{m}}_2 &= -((\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})(\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{m}_1)+(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\boldsymbol{m}_2) \qquad (9)\end{aligned}$$

2.3 Mean of the conditional distribution

It is possible to solve for $\tilde{\boldsymbol{m}}_2$ from (8-9). This yields

$$\begin{aligned}\tilde{\boldsymbol{m}}_2 &= -(\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{(A_1^{-1})}^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2-\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{(A_1^{-1})}^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{21}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2+\boldsymbol{B}_{22})^{-1}\cdot \\ &\quad ((\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})(\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{m}_1)+(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\boldsymbol{m}_2). \qquad (10)\end{aligned}$$

Further, using (4) we can then solve for $\tilde{\boldsymbol{m}}_1$

$$\tilde{\boldsymbol{m}}_1=\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\,\tilde{\boldsymbol{m}}_2. \tag{11}$$

So now we have

$$\tilde{\boldsymbol{m}}=\begin{pmatrix}\tilde{\boldsymbol{m}}_1\\ \tilde{\boldsymbol{m}}_2\end{pmatrix} \tag{12}$$

Then inserting (10) and (11) into (12) we obtain

$$\begin{aligned}\tilde{\boldsymbol{m}} &= \begin{pmatrix}\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{\boldsymbol{C}}_{22}\cdot((\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})(\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{m}_1)+(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\boldsymbol{m}_2)\\ -\tilde{\boldsymbol{C}}_{22}\cdot((\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})(\boldsymbol{A}_1^{-1}\boldsymbol{y}-\boldsymbol{m}_1)+(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\boldsymbol{m}_2)\end{pmatrix} \\ &= \begin{pmatrix}\boldsymbol{A}_1^{-1}+\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{A}_1^{-1}\\ -\tilde{C}_{22}(\boldsymbol{B}_{21}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})k\boldsymbol{A}_1^{-1}\end{pmatrix}\boldsymbol{y} \\ &+ \begin{pmatrix}-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{22}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\\ \tilde{C}_{22}(\boldsymbol{B}_{22}-\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\end{pmatrix}\boldsymbol{m}_1+\begin{pmatrix}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\\ -\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12}-\boldsymbol{B}_{22})\end{pmatrix}\boldsymbol{m}_2\end{aligned} \tag{13}$$

$$\begin{aligned} &= \begin{pmatrix} \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{A}_1^{-1} \\ -\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{A}_1^{-1} \end{pmatrix} \boldsymbol{y} \\ &+ \begin{pmatrix} -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}) & \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{22}) \\ \tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}) & -\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{22}) \end{pmatrix} \boldsymbol{m} \end{aligned} \tag{14}$$

Where by (8) we have

$$\tilde{\boldsymbol{C}}_{22} = (\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2 - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_{\boldsymbol{1}}^{\boldsymbol{-1}})^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{21}\boldsymbol{A}_1^{-1}\boldsymbol{A}_2 + \boldsymbol{B}_{22})^{-1}$$

So the coefficient matrices for $\boldsymbol{y}$ and $\boldsymbol{m}$ that depend on $\boldsymbol{A}$ and $\boldsymbol{B}$ only. We write

$$\tilde{\boldsymbol{m}} = \boldsymbol{G}\boldsymbol{y} + \boldsymbol{H}\boldsymbol{m} \tag{15}$$

where

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{A}_1^{-1} \\ -\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{A}_1^{-1} \end{pmatrix} \tag{16}$$

$$\boldsymbol{H} = \begin{pmatrix} -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}) & \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{22}) \\ \tilde{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{11}) & -\tilde{C}_{22}(\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{22}) \end{pmatrix} \tag{17}$$

There are a lot of common components in the elements of $\boldsymbol{G}$ and $\boldsymbol{H}$. This is more clearly visible if we arrange the terms using the following notation.

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{A}_1^{-1}\boldsymbol{A}_2 && (18) \\ \boldsymbol{J} &= \tilde{\boldsymbol{C}}_{22}(\boldsymbol{B}_{21} - (\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{\mathrm{T}}\boldsymbol{B}_{11}) && (19) \\ &= \tilde{\boldsymbol{C}}_{22}(\boldsymbol{B}_{21} - L^{\mathrm{T}}\boldsymbol{B}_{11}) && (20) \\ \boldsymbol{K} &= \tilde{\boldsymbol{C}}_{22}(\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{\mathrm{T}}\boldsymbol{B}_{12} - B_{22}) && (21) \\ &= \tilde{\boldsymbol{C}}_{22}(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{12} - B_{22}) && (22) \end{aligned}$$

Now $\boldsymbol{G}$, $\boldsymbol{H}$ and $\tilde{\boldsymbol{C}}_{22}$ can be written as

$$\begin{aligned} \boldsymbol{G} &= \begin{pmatrix} -\boldsymbol{A}_1^{-1} + \boldsymbol{L}\boldsymbol{J}\boldsymbol{A}_1^{-1} \\ -\boldsymbol{J}\boldsymbol{A}_1^{-1} \end{pmatrix} && (23) \\ \boldsymbol{H} &= \begin{pmatrix} -\boldsymbol{L}\boldsymbol{J} & \boldsymbol{L}\boldsymbol{K} \\ \boldsymbol{J} & -\boldsymbol{K} \end{pmatrix} && (24) \\ \tilde{\boldsymbol{C}}_{22} &= (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{L} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{21}\boldsymbol{L} + \boldsymbol{B}_{22})^{-1} && (25) \end{aligned}$$

Equation (15) is conditioned on a particular $\boldsymbol{y}$. The result of the iteration is the expected value $E[\tilde{\boldsymbol{m}}]$ over sample of $\boldsymbol{y}$s. If we substitute the relation (3) into (15) it yields

$$\tilde{\boldsymbol{m}} = \boldsymbol{G}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{m}, \tag{26}$$

and the result of the iteration round is the expected value of this, giving

$$\boldsymbol{m}^{(i+1)} = E[\tilde{\boldsymbol{m}}^{(i)}] = \boldsymbol{GA\mu} + \boldsymbol{Hm}^{(i)}. \tag{27}$$

2.4 Covariance matrix of the conditional distribution

2.4.1 The conditional covariance approach

The term (i,j) of covariance matrix of $\boldsymbol{X}$ is by conditional covariace

$$\begin{aligned} \mathrm{Cov}[x_i, x_j] &= E[\mathrm{Cov}[x_i, x_j|\boldsymbol{y}]] + \mathrm{Cov}[E[x_i|\boldsymbol{y}], E[x_j|\boldsymbol{y}]] & (28)\\ &= E[\mathrm{Cov}[x_i, x_j|\boldsymbol{y}]] + \mathrm{Cov}[\tilde{m}_i(y), \tilde{m}_j(y)] & (29)\end{aligned}$$

So the whole covariance matrix can be calculated from

$$\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}] = E[\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}|\boldsymbol{y}]] + \mathrm{Cov}[\tilde{\boldsymbol{m}}(y), \tilde{\boldsymbol{m}}(y)] \tag{30}$$

The latter term, the covariance matrix of $\tilde{\boldsymbol{m}}$ is

$$\begin{aligned} \boldsymbol{S} &= \mathrm{Cov}[\tilde{\boldsymbol{m}}(y), \tilde{\boldsymbol{m}}(y)] & (31)\\ &= E[(\tilde{\boldsymbol{m}}(y) - E[\tilde{\boldsymbol{m}}(y)])(\tilde{\boldsymbol{m}}(y) - E[\tilde{\boldsymbol{m}}(y)])^{\mathrm{T}}] & (32)\\ &= E[(\tilde{\boldsymbol{m}}(y) - \overline{\tilde{\boldsymbol{m}}(y)})(\tilde{\boldsymbol{m}}(y) - \overline{\tilde{\boldsymbol{m}}(y)})^{\mathrm{T}}] & (33)\\ &= E[(\boldsymbol{GAx} - \boldsymbol{GA\mu})(\boldsymbol{GAx} - \boldsymbol{GA\mu})^{\mathrm{T}}] & (34)\\ &= \boldsymbol{GA}\, E[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}](\boldsymbol{GA})^{\mathrm{T}} & (35)\\ &= \boldsymbol{GA\Sigma}(\boldsymbol{GA})^{\mathrm{T}}. & (36)\end{aligned}$$

Where $\boldsymbol{Hm}$ terms from equation (26) cancel out because $\boldsymbol{m}$ as the previous rounds estimate is a constant. Note that the term in the middle is the covariance matrix of $\boldsymbol{Y}$

$$\boldsymbol{\Sigma}_Y = \boldsymbol{A\Sigma A}^{\mathrm{T}}$$

so we do not need to know $\boldsymbol{\Sigma}$ to calculate (36), which would be of course rather inconvenient since that is exactly what we are trying to find out.

Since

$$\boldsymbol{A}_1^{-1}\boldsymbol{A} = \boldsymbol{A}_1^{-1}(\boldsymbol{A}_1, \boldsymbol{A}_2) = (\boldsymbol{I}, \boldsymbol{L}) := \boldsymbol{M} \tag{37}$$

and using the notation $\boldsymbol{M}$

$$\boldsymbol{GA} = \begin{pmatrix} (\boldsymbol{LJ} - \boldsymbol{I})\boldsymbol{M} \\ -\boldsymbol{JM} \end{pmatrix}, \tag{38}$$

and we obtain

$$
\begin{aligned}
\boldsymbol{S} &= \boldsymbol{GA\Sigma}(\boldsymbol{GA})^{\mathrm{T}} && (39)\\
&= \begin{pmatrix} (\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M} \\ -\boldsymbol{JM} \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} ((\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M})^{\mathrm{T}} & (-\boldsymbol{JM})^{\mathrm{T}} \end{pmatrix} && (40)\\
&= \begin{pmatrix} (\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M\Sigma}((\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M})^{\mathrm{T}} & (\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M\Sigma}(-\boldsymbol{JM})^{\mathrm{T}} \\ -\boldsymbol{JM\Sigma}(\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M})^{\mathrm{T}} & \boldsymbol{JM\Sigma}(\boldsymbol{JM})^{\mathrm{T}} \end{pmatrix} && (41)\\
&= \begin{pmatrix} (\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M\Sigma M}^{\mathrm{T}}(L\boldsymbol{J}-\boldsymbol{I})^{\mathrm{T}} & -(\boldsymbol{LJ}-\boldsymbol{I})\boldsymbol{M\Sigma M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \\ -\boldsymbol{JM\Sigma M}^{\mathrm{T}}(\boldsymbol{LJ}-\boldsymbol{I})^{\mathrm{T}} & \boldsymbol{JM\Sigma M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \end{pmatrix} && (42)
\end{aligned}
$$

This is the expected value of sample covariance matrix for $\tilde{\boldsymbol{m}}$ and also corresponds to one produced by an infinite sample of measurements $\boldsymbol{y}$. Note that the $n \times n$ matrix $\boldsymbol{GA}$ has rank that is at most the column rank of $\boldsymbol{A}$ which is m. Therefore it is singular, and hence $\boldsymbol{S}$ is singular.

The first term $\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}|\boldsymbol{y}]$, the conditional covariance matrix for $\boldsymbol{x}$ conditioned on $\boldsymbol{y}$ can be obtained by calculating

$$
\begin{aligned}
\tilde{\boldsymbol{C}}_{11} &= E[(\boldsymbol{x}_1 - \tilde{\boldsymbol{m}}_1)(\boldsymbol{x}_1 - \tilde{\boldsymbol{m}}_1)^{\mathrm{T}}]\\
&= E[(\boldsymbol{A}_1^{-1}\boldsymbol{y} - \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{x}_2 - \boldsymbol{A}_1^{-1}\boldsymbol{y} + \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{\boldsymbol{m}}_2)\cdot\\
&\quad (\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}} - \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}} - \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}} + \tilde{\boldsymbol{m}}_2^{\mathrm{T}}\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}})]\\
&= E[(-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2))(-(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}})\\
&= \boldsymbol{A}_1^{-1}\boldsymbol{A}_2 E[(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)((\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}]\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\\
&= \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{\boldsymbol{C}}_{22}\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}\\
&= \boldsymbol{L}\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} && (43)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\boldsymbol{C}}_{12} &= E[(\boldsymbol{x}_1 - \tilde{\boldsymbol{m}}_1)(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}]\\
&= E[(\boldsymbol{A}_1^{-1}\boldsymbol{y} - \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{x}_2 - \boldsymbol{A}_1^{-1}\boldsymbol{y} + \boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{\boldsymbol{m}}_2)(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}\\
&= E[(-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2))(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}]\\
&= -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2 E[(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}]\\
&= -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\tilde{\boldsymbol{C}}_{22}\\
&= -\boldsymbol{L}\tilde{\boldsymbol{C}}_{22} && (44)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\boldsymbol{C}}_{21} &= E[(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)(\boldsymbol{x}_1 - \tilde{\boldsymbol{m}}_1)^{\mathrm{T}}]\\
&= -\tilde{\boldsymbol{C}}_{22}\boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{A}_1^{-1})^{\mathrm{T}}
\end{aligned}
$$

$$= -\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} \tag{45}$$

$$\begin{aligned}
\tilde{\boldsymbol{C}}_{22} &= E[(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)(\boldsymbol{x}_2 - \tilde{\boldsymbol{m}}_2)^{\mathrm{T}}] \\
&= (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{L} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{21}\boldsymbol{L} + \boldsymbol{B}_{22})^{-1} \qquad (46)
\end{aligned}$$

$$\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}|\boldsymbol{y}] = \tilde{\boldsymbol{C}} = \begin{pmatrix} \tilde{\boldsymbol{C}}_{11} & \tilde{\boldsymbol{C}}_{12} \\ \tilde{\boldsymbol{C}}_{21} & \tilde{\boldsymbol{C}}_{22} \end{pmatrix}, \tag{47}$$

Since $\tilde{\boldsymbol{C}}$ in (47) depends on $\boldsymbol{A}$ and $\boldsymbol{B}$ only, that is not on $\boldsymbol{X}$, it is constant with regard to the expectation operation. Hence

$$E[\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}|\boldsymbol{y}]] = E[\tilde{\boldsymbol{C}}] = \tilde{\boldsymbol{C}} \tag{48}$$

So now the updated covariance matrix in (30) can be written as

$$\begin{aligned}
\boldsymbol{C}^{(i+1)} &= \tilde{\boldsymbol{C}} + \boldsymbol{S} \qquad (49) \\
&= \tilde{\boldsymbol{C}} + \boldsymbol{G}\boldsymbol{A}\boldsymbol{\Sigma}(\boldsymbol{G}\boldsymbol{A})^{\mathrm{T}} \qquad (50) \\
&= \begin{pmatrix} \boldsymbol{L}\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} & -\boldsymbol{L}\tilde{\boldsymbol{C}}_{22} \\ -\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} & \tilde{\boldsymbol{C}}_{22} \end{pmatrix} \\
&+ \begin{pmatrix} (\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})^{\mathrm{T}} & -(\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \\ -\boldsymbol{J}\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})^{\mathrm{T}} & \boldsymbol{J}\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \end{pmatrix} \qquad (51) \\
&= \begin{pmatrix} \boldsymbol{L}\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} + (\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})^{\mathrm{T}} & -\boldsymbol{L}\tilde{\boldsymbol{C}}_{22} - (\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \\ -\tilde{\boldsymbol{C}}_{22}\boldsymbol{L}^{\mathrm{T}} - \boldsymbol{J}\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{J}-\boldsymbol{I})^{\mathrm{T}} & \tilde{\boldsymbol{C}}_{22} + \boldsymbol{J}\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}} \end{pmatrix} \\
& \qquad (52)
\end{aligned}$$

Where $\boldsymbol{L}$ and $\boldsymbol{M}$ depend only on $\boldsymbol{A}$ and thus, along with the real covariance matrix $\boldsymbol{\Sigma}$, do not change during iteration. On the other hand $\tilde{\boldsymbol{C}}_{22}$ and $\boldsymbol{J}$ are dependent on $\boldsymbol{B}$ and thus do change.

2.4.2 The co-ordinate transformation approach

The conditional covariance matrix for $\boldsymbol{x}$ was calculated in equations (43)-(47) as

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} \tilde{\boldsymbol{C}}_{11} & \tilde{\boldsymbol{C}}_{12} \\ \tilde{\boldsymbol{C}}_{21} & \tilde{\boldsymbol{C}}_{22} \end{pmatrix}, \tag{53}$$

Because of condition (3) there cannot be any variance in the direction of $\boldsymbol{y}$ vectors. All the variance is orthogonal to these directions, as is shown in figure 3 for the two dimensional case. So transformed co-ordinates can be selected such that there is variation only in $n - m$ directions. First m axes are the $\boldsymbol{y}$ vectors

$$\boldsymbol{y}_i = \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{A}$$

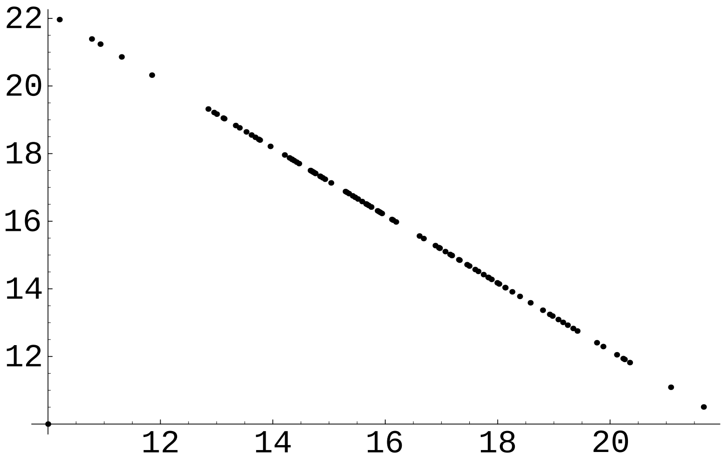

Figure 3: Sample conditioned on $\boldsymbol{y}$ with 100 points generated from the conditional sample.

where $\boldsymbol{e}_i$ is an m-vector, where the ith element is one and all the others are zero. This way we can obtain m directions of the new co-ordinate axes. Let us denote the rest of the co-ordinate axes by $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-m})$. The z_i-vectors have to be orthogonal to these and also be selected so that there are no covariance terms in the transformed covariance matrix.

So let us make the transformation from X to X', such that x'_i are the unit vectors of y_i and z_i, and $\boldsymbol{T}$ is the transformation matrix compiled from their unit vectors. Then mean and covariance matrix of $\boldsymbol{x}$ and the real covariance matrix are

$$\begin{aligned} \boldsymbol{m}' &= \boldsymbol{T}\boldsymbol{m} && (54) \\ \tilde{\boldsymbol{C}}' &= \boldsymbol{T}\tilde{\boldsymbol{C}}\boldsymbol{T}^{\mathrm{T}} && (55) \\ \boldsymbol{\Sigma}' &= \boldsymbol{T}\boldsymbol{\Sigma}\boldsymbol{T}^{\mathrm{T}} && (56) \end{aligned}$$

Now we could draw a random sample with mean and variance as above, so that the values $x'_1, \ldots, x'_m$ remain constant in the sample, and $x'_{m+1}, \ldots, x'_{m+n}$ vary. **Note that since we hold the first m values constant, it follows that this approach is valid only when we have a prior where all the variances are equal.** Thus we get a random sample of the conditional distribution, then return these to the original co-ordinates. Then we can pick the next measurement $\boldsymbol{y}$ and repeat the procedure. This way we get the distribution conditioned on $\boldsymbol{y}$ with several different $\boldsymbol{y}$. This method is illustrated in figure 4.

The expectation of the Covariance matrix of the sample cloud can be obtained by

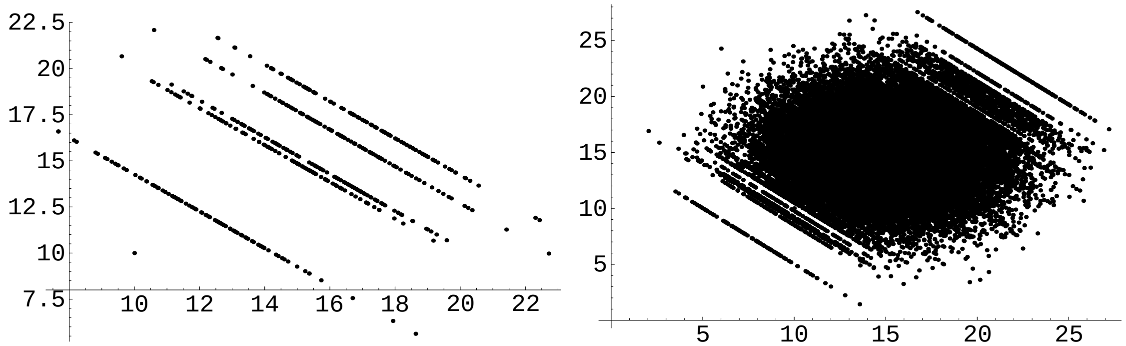

Figure 4: Conditional sample. On the left with five different $\boldsymbol{y}$ values, on the right with 500 values.

selecting the the non-zero term from $\tilde{C}'$ and the rest of the terms from Σ', so that covariance between these terms is zero. Then returning this matrix $C'^{(1)}$ to original co-ordinates

$$C^{(1)} = T^{-1}C'^{(1)}(T^{-1})^{\mathrm{T}} \tag{57}$$

An example of this is given in Section 3.4.

3 Example: Network with one link

3.1 Introduction

To illustrate the behavior of the method we consider an example that is simplified to extremity. Consider the network in figure 5. It has a single link and two traffic flows.

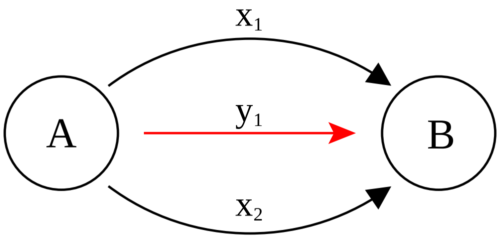

Figure 5: Two dimensional example network

This is an unrealistic example, since both x_1 and x_2 should belong to OD pair AB and there are no way to differentiate how much a specific x_i contributes to the link load $\boldsymbol{y}$. However, because of this, the solution depends heavily on the prior estimate, and thus gives good opportunity to study the effect of different priors, as well as display graphically the situation. In this section we derive the same results for this example network as in section 2 for the general case.

3.2 Mean

Let us first define

$$\begin{aligned} \boldsymbol{A} &= \begin{pmatrix} 1 & 1 \end{pmatrix} && (58)\\ \boldsymbol{\Sigma} &= \begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix} && (59)\\ \boldsymbol{\Sigma}_Y &= \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^{\mathrm{T}} = r + s && (60)\\ \boldsymbol{B} &= \boldsymbol{C}^{-1} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} && (61) \end{aligned}$$

And, using notation w for simplicity

$$\begin{aligned} w^{-1} = \tilde{\boldsymbol{C}}_{22} &= (b_{11} - b_{12} - b_{21} + b_{22})^{-1} && (62)\\ \tilde{\boldsymbol{C}} &= \begin{pmatrix} w^{-1} & -w^{-1} \\ -w^{-1} & w^{-1} \end{pmatrix} && (63) \end{aligned}$$

equations (16)-(17) for $\boldsymbol{G}$ and $\boldsymbol{H}$ give

$$\begin{aligned} \boldsymbol{G} &= \begin{pmatrix} 1 - \frac{(b_{11}-b_{21})}{w} \\ \\ \frac{(b_{11}-b_{21})}{w} \end{pmatrix} = \begin{pmatrix} \frac{(b_{22}-b_{12})}{w} \\ \\ \frac{(b_{11}-b_{21})}{w} \end{pmatrix} && (64)\\ \boldsymbol{H} &= \begin{pmatrix} \frac{b_{11}-b_{21}}{w} & \frac{b_{12}-b_{22}}{w} \\ \\ \frac{b_{21}-b_{11}}{w} & \frac{b_{22}-b_{12}}{w} \end{pmatrix} && (65) \end{aligned}$$

Now we can calculate the estimate for mean using equation (81).

$$\begin{aligned} E[\tilde{\boldsymbol{m}}] &= \boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{m} && (66)\\ &= \begin{pmatrix} \frac{(b_{22}-b_{12})(\mu_1+\mu_2)}{w} \\ \frac{(b_{11}-b_{21})(\mu_1+\mu_2)}{w} \end{pmatrix} + \begin{pmatrix} \frac{(b_{11}-b_{21})m_1+(b_{12}-b_{22})m_2}{w} \\ \frac{(b_{21}-b_{11})m_1+(b_{22}-b_{12})m_2}{w} \end{pmatrix} && (67)\\ &= \frac{1}{w}\begin{pmatrix} (b_{11}-b_{21})m_1 + (b_{22}-b_{12})((\mu_1+\mu_2)-m_2) \\ (b_{11}-b_{21})((\mu_1+\mu_2)-m_1) + (b_{22}-b_{12})m_2 \end{pmatrix} && (68) \end{aligned}$$

It follows from the equation that the result for $E[\tilde{\boldsymbol{m}}]$ is found by moving to the direction specified by the variances of the prior distribution. The narrower the distribution the more certain we are of this prior, and the movement is larger in the other direction where the uncertainty is larger. This is shown in figure (6).

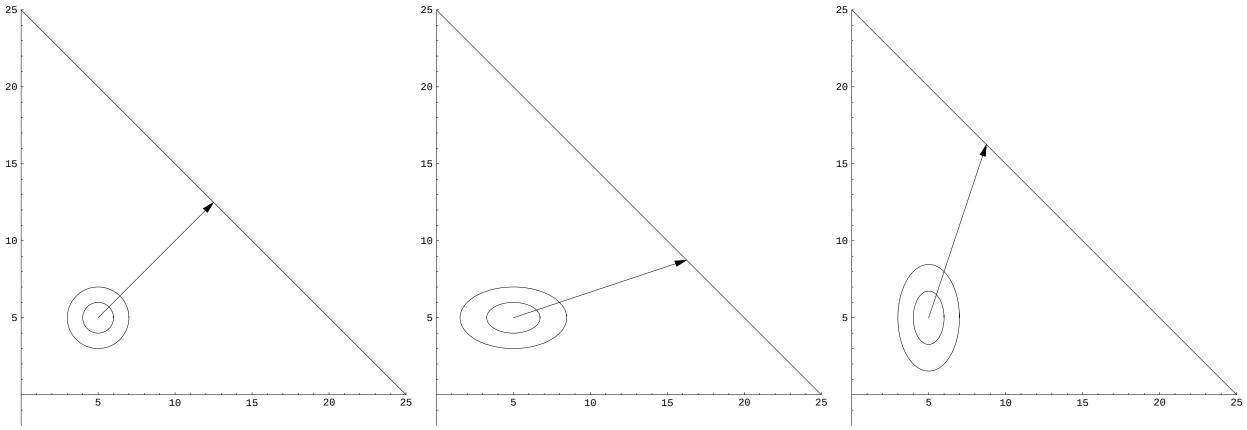

Figure 6: Covariance matrix of the prior distribution affects the mean of the result distribution.

3.3 Covariance matrix

Equation (36) for $\boldsymbol{S}$ now yields

$$\boldsymbol{S} = \boldsymbol{G}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}} = \begin{pmatrix} \frac{(b_{22}-b_{12})^2(r+s)}{w^2} & \frac{(b_{11}-b_{21})(b_{22}-b_{12})(r+s)}{w^2} \\ \\ \frac{((b_{11}-b_{21})(b_{22}-b_{12})(r+s)}{w^2} & \frac{((b_{11}-b_{21})^2(r+s))}{w^2} \end{pmatrix} \tag{69}$$

and $\tilde{\boldsymbol{C}}$ is as given in (63), so covariance matrix is

$$\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}] = \tilde{\boldsymbol{C}} + \boldsymbol{S} \tag{70}$$

$$= \begin{pmatrix} w^{-1} & -w^{-1} \\ -w^{-1} & w^{-1} \end{pmatrix} + \begin{pmatrix} \frac{(b_{22}-b_{12})^2(r+s)}{w^2} & \frac{(b_{11}-b_{21})(b_{22}-b_{12})(r+s)}{w^2} \\ \frac{(b_{11}-b_{21})(b_{22}-b_{12})(r+s)}{w^2} & \frac{(b_{11}-b_{21})^2(r+s)}{w^2} \end{pmatrix} \tag{71}$$

$$= \frac{1}{w^2} \begin{pmatrix} (b_{22}-b_{12})^2(r+s)+w & (b_{11}-b_{21})(b_{22}-b_{12})(r+s)-w \\ (b_{11}-b_{21})(b_{22}-b_{12})(r+s)-w & (b_{11}-b_{21})^2(r+s)+w \end{pmatrix} \tag{72}$$

3.4 Covariance matrix by transformation method

First we have need obtain the transformation matrix T. The direction of no variance in two dimensional case is $y_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}$ which is the only row of routing matrix A. The direction of free variability is then orthogonal to this, that is $z_1 = \begin{pmatrix} -1 & 1 \end{pmatrix}$. Normalizing these we obtain

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{73}$$

Then from (55)-(56)

$$\tilde{\boldsymbol{C}}' = \begin{pmatrix} 2w^{-1} & 0 \\ 0 & 0 \end{pmatrix} \tag{74}$$

$$\boldsymbol{\Sigma}' = \frac{1}{2}\begin{pmatrix} r+s & r-s \\ r-s & r+s \end{pmatrix} \tag{75}$$

We pick the non zero term from C' and from Σ' the terms that are not in the same row or column with the term already picked, since all covariance including the C' term are set to zero.

$$\tilde{\boldsymbol{C}}'^{(1)} = \begin{pmatrix} 2w^{-1} & 0 \\ 0 & \frac{1}{2}(r+s) \end{pmatrix} \tag{76}$$

And finally we use equation (57) to return this to original co-ordinates.

$$C^{(1)} = \begin{pmatrix} \frac{1}{4}(r+s)+w^{-1} & \frac{1}{4}(r+s)-w^{-1} \\ \frac{1}{4}(r+s)-w^{-1} & \frac{1}{4}(r+s)+w^{-1} \end{pmatrix} \tag{77}$$

The transformation method here is defined so that it gives correct answers only when prior variances of x_1 and x_2 are the same. So we have to set $b_{11} = b_{22} = b$ and $b_{12} = b_{21} = 0$ in (72). So now

$$\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}] = \frac{1}{w^2}\begin{pmatrix} b^2(r+s)+w & b^2(r+s)-w \\ b^2(r+s)-w & b^2(r+s)+w \end{pmatrix} \tag{78}$$

$$= \begin{pmatrix} \frac{b^2}{(2b)^2}(r+s)+\frac{w}{w^2} & \frac{b^2}{(2b)^2}(r+s)-\frac{w}{w^2} \\ \frac{b^2}{(2b)^2}(r+s)-\frac{w}{w^2} & \frac{b^2}{(2b)^2}(r+s)+\frac{w}{w^2} \end{pmatrix} \tag{79}$$

$$= \begin{pmatrix} \frac{1}{4}(r+s)+w^{-1} & \frac{1}{4}(r+s)-w^{-1} \\ \frac{1}{4}(r+s)-w^{-1} & \frac{1}{4}(r+s)+w^{-1} \end{pmatrix} \tag{80}$$

Where we used the fact that $w = (b_{11} - b_{12} - b_{21} + b_{22}) = 2b$.

And we can see that (72) and (77) indeed yield the same result, as we would expect.

4 Iteration

4.1 Introduction

In this section the equations for the iteration are given. Then we show that in this situation where we assume a single underlying distribution, the iteration is useless as the result converges to its final values after the first step. This is proven for the mean in general case in theorem 1. Examples show that the same is true for the covariance matrix in two-dimensional case, and in fact for all the examples we have studied. The proof for the general case is left for future work.

4.2 General case

The expectation of the result of an iteration round is

$$\begin{cases} \boldsymbol{m}^{(i+1)} &= \boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{m}^{(i)} \\ \boldsymbol{C}^{(i+1)} &= \tilde{\boldsymbol{C}} + \boldsymbol{G}\boldsymbol{A}\boldsymbol{\Sigma}(\boldsymbol{G}\boldsymbol{A})^{\mathrm{T}} \end{cases} \tag{81}$$

where $\boldsymbol{G}$, $\boldsymbol{H}$ and $\tilde{\boldsymbol{C}}$ depend on $\boldsymbol{C}^{(i)}$ through $\boldsymbol{B}^{(i)}$, while $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$ remain constant.

Theorem 1 The expected value of the mean $\boldsymbol{m}$ does not change in the iteration after the first iteration. That is

$$\boldsymbol{m}^{(i+1)} = \boldsymbol{m}^{(i)} \qquad \forall i > 1. \tag{82}$$

Proof From (27) we can see that

$$\begin{aligned} \boldsymbol{m}^{(i)} &= \boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{m}^{(i-1)} && (83) \\ \boldsymbol{m}^{(i+1)} &= \boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}(\boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{m}^{(i-1)}) && (84) \\ \boldsymbol{m}^{(i+1)} - \boldsymbol{m}^{(i)} &= \boldsymbol{H}(\boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{m}^{(i-1)}) - \boldsymbol{H}\boldsymbol{m}^{(i-1)} && (85) \\ &= \boldsymbol{H}\boldsymbol{G}\boldsymbol{A}\boldsymbol{\mu} + (\boldsymbol{H}\boldsymbol{H} - \boldsymbol{H})\boldsymbol{m}^{(i-1)} && (86) \end{aligned}$$

So to prove that $\boldsymbol{m}^{(i+1)} = \boldsymbol{m}^{(i)}$ we need to show that

$$\begin{aligned} \boldsymbol{H}\boldsymbol{G} &= \boldsymbol{0} && (87) \\ \boldsymbol{H}\boldsymbol{H} &= \boldsymbol{H} && (88) \end{aligned}$$

First let us show the following

$$\begin{aligned} \boldsymbol{J}\boldsymbol{L} + \boldsymbol{K} &= \boldsymbol{C}_{22}(\boldsymbol{B}_{21} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{11})\boldsymbol{L} + \boldsymbol{C}_{22}(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{22}) && (89) \\ &= -\boldsymbol{C}_{22}(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{11}\boldsymbol{L} - \boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}_{12} - \boldsymbol{B}_{21}\boldsymbol{L} + \boldsymbol{B}_{22}) && (90) \\ &= -\boldsymbol{C}_{22}(\boldsymbol{C}_{22})^{-1} && (91) \\ &= -\boldsymbol{I} && (92) \end{aligned}$$

and we can use this result to proof the relations

$$\begin{aligned}
\boldsymbol{HG} &= \begin{pmatrix} -\boldsymbol{LJA}_1^{-1} - \boldsymbol{LJLJA}_1^{-1} - \boldsymbol{LKJA}_1^{-1} \\ \boldsymbol{JA}_1^{-1} + \boldsymbol{JLJA}_1^{-1} + \boldsymbol{KJA}_1^{-1} \end{pmatrix} && (93) \\
&= \begin{pmatrix} -\boldsymbol{L}(\boldsymbol{I} + \boldsymbol{JL} + \boldsymbol{K})\boldsymbol{JA}_1^{-1} \\ (\boldsymbol{I} + \boldsymbol{JL} + \boldsymbol{K})\boldsymbol{JA}_1^{-1} \end{pmatrix} && (94) \\
&= \begin{pmatrix} -\boldsymbol{L}(\boldsymbol{I} - \boldsymbol{I})\boldsymbol{JA}_1^{-1} \\ (\boldsymbol{I} - \boldsymbol{I})\boldsymbol{JA}_1^{-1} \end{pmatrix} && (95) \\
&= \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} && (96)
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{HH} &= \begin{pmatrix} \boldsymbol{LJLJ} + \boldsymbol{LKJ} & -\boldsymbol{LJLK} - \boldsymbol{LKK} \\ -\boldsymbol{JLJ} - \boldsymbol{KJ} & \boldsymbol{JLK} + \boldsymbol{KK} \end{pmatrix} && (97) \\
&= \begin{pmatrix} -\boldsymbol{L}(-\boldsymbol{JL} - \boldsymbol{K})\boldsymbol{J} & \boldsymbol{L}(-\boldsymbol{JL} - \boldsymbol{K})\boldsymbol{K} \\ (-\boldsymbol{JL} - \boldsymbol{K})\boldsymbol{J} & -(-\boldsymbol{JL} - \boldsymbol{K})\boldsymbol{K} \end{pmatrix} && (98) \\
&= \begin{pmatrix} -\boldsymbol{L}(\boldsymbol{I})\boldsymbol{J} & \boldsymbol{L}(\boldsymbol{I})\boldsymbol{K} \\ (\boldsymbol{I})\boldsymbol{J} & -(\boldsymbol{I})\boldsymbol{K} \end{pmatrix} && (99) \\
&= \begin{pmatrix} -\boldsymbol{LJ} & \boldsymbol{LK} \\ \boldsymbol{J} & -\boldsymbol{K} \end{pmatrix} = \boldsymbol{H} && (100)
\end{aligned}$$

So

$$\begin{aligned}
\boldsymbol{m}^{(i+1)} - \boldsymbol{m}^{(i)} &= \boldsymbol{HGA\mu} + (\boldsymbol{HH} - \boldsymbol{H})\boldsymbol{m}^{(i-1)} && (101) \\
&= \mathbf{0} \cdot \boldsymbol{A\mu} + (\boldsymbol{H} - \boldsymbol{H})\boldsymbol{m}^{(i-1)} && (102) \\
&= \mathbf{0} && (103)
\end{aligned}$$

And this completes the proof.

Theorem 2 The ecpexted value of the estimated covariance matrix $\boldsymbol{C}$ does not change in the iteration after the first iteration. That is

$$\boldsymbol{C}^{(i+1)} = \boldsymbol{C}^{(i)} \qquad \forall i > 1. \tag{104}$$

The proof has not been completed for the general case. In the following section it is proven for the two dimensional case. All numerical examples of various topologies indicate that the covariance matrix also converges after the first iteration step.

4.3 Two dimensional case

4.3.1 Mean

In Theorem 1 it is shown that the result does not change after first iteration. The proof is given in matrix form. Here the same result is proven as an example specifically for two dimensional case.

After first iteration the estimate $\boldsymbol{m}^{(1)} = (\, m_1 \quad m_2 \,)$ is on the line that goes through points $(\, \mu_1 \quad \mu_2 \,)$, $(\, \mu_1 + \mu_2 \quad 0 \,)$ and $(\, 0 \quad \mu_1 + \mu_2 \,)$ as can be seen from figure 8. It obviously satisfies

$$\begin{aligned} \boldsymbol{A}\boldsymbol{m} &= \boldsymbol{A}\boldsymbol{\mu} && (105) \\ m_1 + m_2 &= \mu_1 + \mu_2 && (106) \end{aligned}$$

and can thus be written

$$\boldsymbol{m}^{(1)} = \begin{pmatrix} m1 \\ (\mu_1 + \mu_2) - m1 \end{pmatrix}. \tag{107}$$

Then from (68)

$$\begin{aligned} \boldsymbol{m}^{(2)} &= \frac{1}{w} \begin{pmatrix} (b_{11} - b_{21})m_1 + (b_{22} - b_{12})((\mu_1 + \mu_2) - m_2) \\ (b_{11} - b_{21})((\mu_1 + \mu_2) - m_1) + (b_{22} - b_{12})m_2 \end{pmatrix} && (108) \\ &= \frac{1}{w} \begin{pmatrix} (b_{11} - b_{21})m_1 + (b_{22} - b_{12})((\mu_1 + \mu_2) - ((\mu_1 + \mu_2) - m_1) \\ (b_{11} - b_{21})((\mu_1 + \mu_2) - m_1) + (b_{22} - b_{12})((\mu_1 + \mu_2) - m_1) \end{pmatrix} && (109) \\ &= \frac{1}{w} \begin{pmatrix} (b_{11} - b_{21} - b_{12} + b_{12})(m_1) \\ (b_{11} - b_{21} - b_{12} + b_{22})((\mu_1 + \mu_2) - m_1) \end{pmatrix} && (110) \\ &= \begin{pmatrix} m_1 \\ (\mu_1 + \mu_2) - m_1 \end{pmatrix} = \boldsymbol{m}^{(1)} && (111) \end{aligned}$$

4.3.2 Covariance matrix

The sum of the elements of the Covariance matrix $\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}]$ after first iteration is $r + s$, the sum of the real variances of OD pair distributions. This can be seen from the fact that the coefficients of $(r + s)$ in elements of (69) sum to one, and the elements of (63) obviously sum to zero.

$$\boldsymbol{C}^{(1)} = \mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}^{\mathrm{T}}] = \begin{pmatrix} v_1 & c_{12} \\ c_{12} & v_2 \end{pmatrix} \tag{112}$$

where $v_1 + v_2 + 2c_{12} = (r+s)$.

$$\begin{aligned} \boldsymbol{B} &= (\boldsymbol{C}^{(1)})^{-1} = \frac{1}{v_1 v_2 - c_{12}^2} \begin{pmatrix} v_2 & -c_{12} \\ -c_{12} & v_1 \end{pmatrix} && (113) \\ w &= b_{11} - b_{12} - b_{21} + b_{22} = \frac{v_1 + v_2 + 2c_{12}}{v_1 v_2 - c_{12}^2} && (114) \end{aligned}$$

Let us look at the upper left element of (72) in the second iteration

$$\begin{aligned} (\boldsymbol{C}^{(2)})_{11} &= \frac{(b_{22} - b_{12})^2}{w^2}(r+s) + w^{-1} && (115) \\ &= \frac{\frac{(v_1+c_{12})^2}{(v_1v_2-c_{12}^2)^2}}{\left(\frac{v_1+v_2+2c_{12}}{v_1v_2-c_{12}^2}\right)^2}(r+s) + \frac{v_1v_2 - c_{12}^2}{v_1+v_2+2c_{12}} && (116) \\ &= \frac{(v_1+c_{12})^2}{(v_1+v_2+2c_{12})^2}(v_1+v_2+2c_{12}) + \frac{v_1v_2 - c_{12}^2}{v_1+v_2+2c_{12}} && (117) \\ &= \frac{v_1^2 + 2v_1c_{12} + c_{12}^2 + v_1v_2 - c_{12}^2}{v_1+v_2+2c_{12}} && (118) \\ &= \frac{v_1(v_1 + 2c_{12} + v_2)}{v_1+v_2+2c_{12}} && (119) \\ &= v_1 && (120) \end{aligned}$$

And similar results for other elements can be obtained to show that

$$\boldsymbol{C}^{(2)} = \begin{pmatrix} v_1 & c_{12} \\ c_{12} & v_2 \end{pmatrix} = \boldsymbol{C}^{(1)} \qquad (121)$$

5 Numerical examples

5.1 Introduction

In this section specific numerical examples are discussed to show the behavior of the method in different situations. First the two dimensional example of section 3 is studied. Then a more realistic, yet still very small, example network with two links and three separate OD pairs is considered. The topology for this network is shown in figure 7.

5.2 Two dimensional numerical example

Choosing the values

$$\boldsymbol{m} = \begin{pmatrix} 2 & 4 \end{pmatrix} \qquad (122)$$

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{123}$$

$$B = C^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{124}$$

$$\mu = \begin{pmatrix} 15 & 10 \end{pmatrix} \tag{125}$$

$$\Sigma = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} \tag{126}$$

Inserting these values to equation (81) (or to the equations specific to the two dimensional case (68) and (72)) yields

$$\tilde{m}^{(i+1)} = \begin{pmatrix} 11.5 & 13.5 \end{pmatrix} \tag{127}$$

$$C^{(i+1)} = \begin{pmatrix} 3.75 & 2.75 \\ 2.75 & 3.75 \end{pmatrix} \tag{128}$$

The results are illustrated in figure 8. In figure 9 the sample cloud of the transformation method is plotted in the same picture.

5.3 Three dimensional numerical example

example1

$$\mu = \begin{pmatrix} 5 & 10 & 15 \end{pmatrix} \tag{129}$$

$$\Sigma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \tag{130}$$

$$m^{(0)} = \begin{pmatrix} 5 & 5 & 5 \end{pmatrix} \tag{131}$$

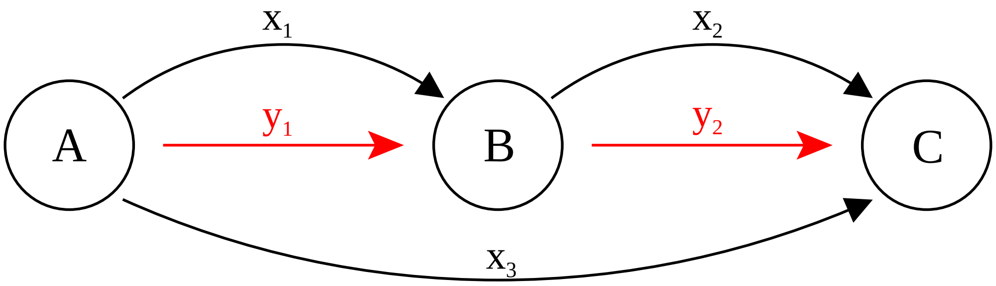

Figure 7: Example network considered in section 5.3

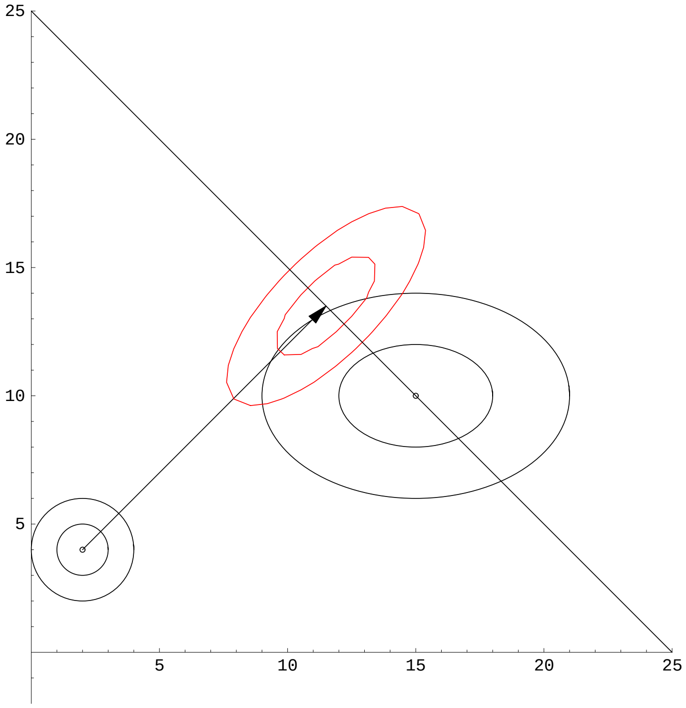

Figure 8: Result of the iteration. The arrow shows the movement of the estimate from prior distribution to new estimate distribution shown in red. The real distribution $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is shown in black

$$\boldsymbol{C}^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{132}$$

With these starting values the method yields the following result

$$\boldsymbol{m}^{(1)} = \begin{pmatrix} 6.67 & 11.67 & 13.33 \end{pmatrix} \tag{133}$$

$$\boldsymbol{C}^{(1)} = \begin{pmatrix} 3.33 & -1.00 & 1.33 \\ -1.00 & 2.67 & 0.67 \\ 1.33 & 0.67 & 3.00 \end{pmatrix} \tag{134}$$

In figure 10 the result is shown as two dimensional projection for each co-ordinate plane. In x_1, x_3-plane and x_2, x_3-plane the line where the new mean estimate is located comes from equation

$$\boldsymbol{A\mu} = \boldsymbol{Ax} \tag{135}$$

which in three-dimensional case becomes

$$\mu_1 + \mu_3 = x_1 + x_3 \tag{136}$$

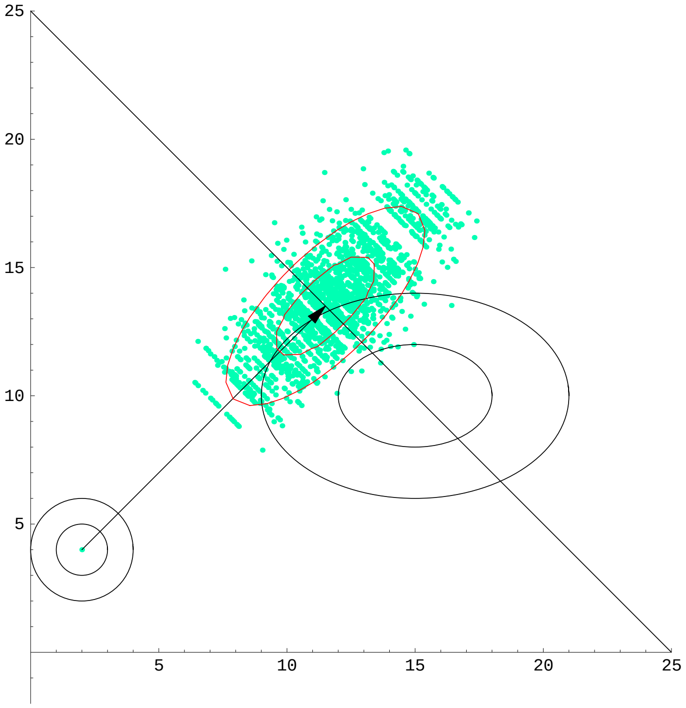

Figure 9: Sample cloud produced by co-ordinate transformation method coincides with explicit result obtained by conditional covariance method.

$$\mu_2 + \mu_3 \quad = \quad x_2 + x_3 \tag{137}$$

and are equations for planes in three dimensional space. The line in x_1, x_2-plane is the transversal of these planes

$$x_1 - x_2 \quad = \quad \mu_1 - \mu_2 \tag{138}$$

example2

Let us change the prior estimate for the covariance matrix.

$$\boldsymbol{\mu} \quad = \quad \begin{pmatrix} 5 & 10 & 15 \end{pmatrix} \tag{139}$$

$$\boldsymbol{\Sigma} \quad = \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \tag{140}$$

$$\boldsymbol{m}^{(0)} \quad = \quad \begin{pmatrix} 5 & 5 & 5 \end{pmatrix} \tag{141}$$

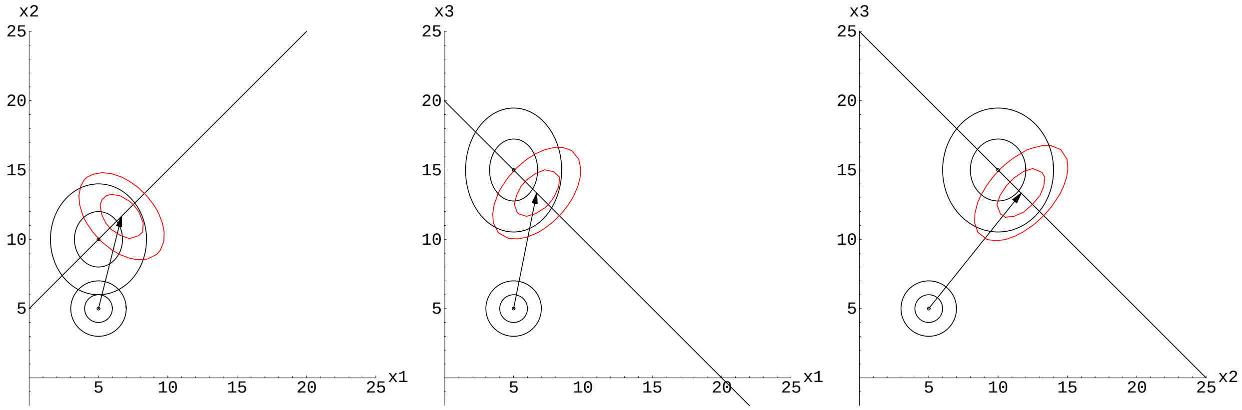

Figure 10: Example 1

$$\boldsymbol{C}^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{142}$$

The result is now

$$\boldsymbol{m}^{(1)} = \begin{pmatrix} 6.82 & 11.82 & 13.18 \end{pmatrix} \tag{143}$$

$$\boldsymbol{C}^{(1)} = \begin{pmatrix} 1.67 & -0.42 & 1.06 \\ -0.42 & 4.48 & 0.15 \\ 1.06 & 0.15 & 4.21 \end{pmatrix} \tag{144}$$

And is shown in figure 11. Notice that the estimate for the mean is very close to that in example 1, even though we changed the covariance prior. This is because the difference between m_1 and m_2 is fixed, because it is known from the difference between the measured link loads. The results move together in the iteration so uncertainty in the form of large prior variance for one of them does not affect the result as clearly as was the case in figure 6 for two-dimensions, where no limiting relations are known between elements of $\boldsymbol{m}$.

example3

$$\boldsymbol{\mu} = \begin{pmatrix} 5 & 10 & 15 \end{pmatrix} \tag{145}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \tag{146}$$

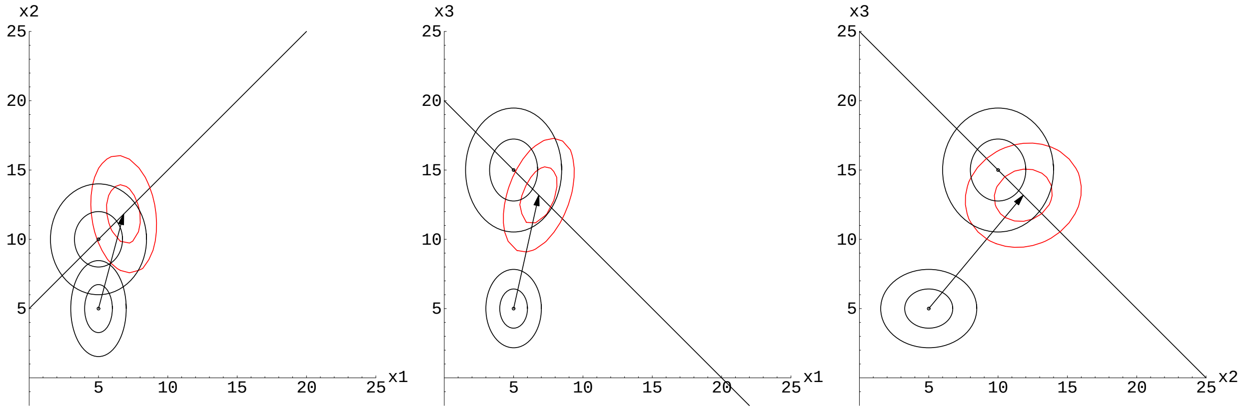

Figure 11: Example 2

$$\boldsymbol{m}^{(0)} = \begin{pmatrix} 5 & 5 & 5 \end{pmatrix} \tag{147}$$

$$\boldsymbol{C}^{(0)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{148}$$

Now we have changed the covariance prior so that the difference between variances of x_2 and x_3 is large. There is no relation known between them, so no the change affects the result more dramatically. This is shown in figure 12.

$$\boldsymbol{m}^{(1)} = \begin{pmatrix} 10 & 15 & 10 \end{pmatrix} \tag{149}$$

$$\boldsymbol{C}^{(1)} = \begin{pmatrix} 3.82 & 1.10 & 1.18 \\ 1.10 & 5.38 & 0.89 \\ 1.18 & 0.89 & 1.82 \end{pmatrix} \tag{150}$$

6 Conclusion

6.1 Summary

In this report a case study on traffic matrix estimation under Gaussian OD pair traffic distributions was made. We studied the behavior of a special case of the Vaton-Gravey method inferring estimates for OD pair traffic demands based on link counts and some prior distribution. In Section 2 the expected values for parameters of the conditional distribution of $\boldsymbol{x}$ conditioned on $\boldsymbol{y}$ were computed by utilizing the characteristics of

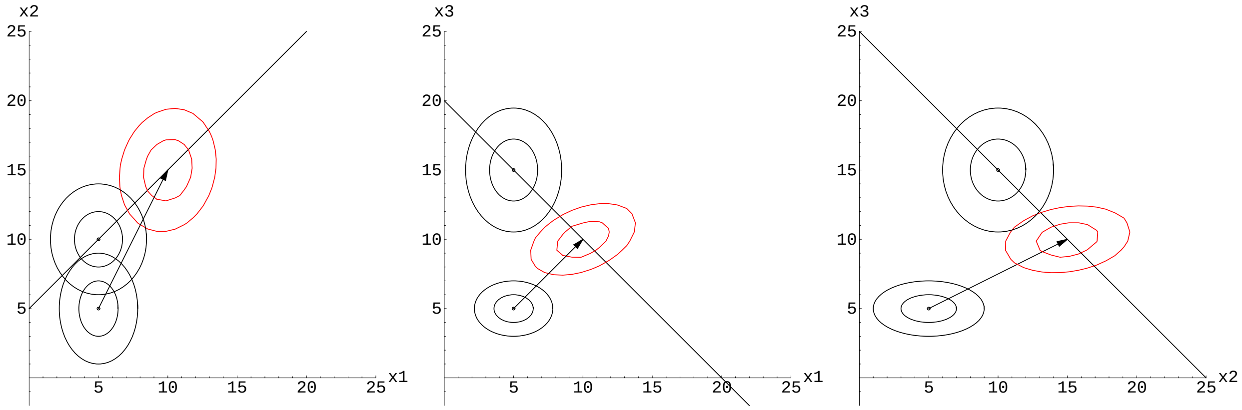

Figure 12: Example 3

the normal distribution. By writing the density function of the conditional distribution in the complete square form, we were able to pick out necessary terms to obtain expressions for conditional mean and covariance estimators. Then taking expectation of these over the link counts y, we obtained the estimators for the distribution of OD pairs. This method was illustrated through simple examples in Section 3.

The main result of the report was given in Section 4 were we state that the iteration converges after first step, and proved this result for the estimator of the mean. However, the proof for the covariance matrix estimator has not yet been completed. This is left as future work. The result indicates that under the assumptions we make about the distributions there is no benefit of iteration. The first estimator is as accurate as it can come in this situation.

In Section 5 some numerical examples were studied in simple topologies. In the degenerate case of one link and two OD pairs the prior chosen determined largely the result, since link count data obviously cannot give any indication about the OD pair traffic amounts. In realistic, yet still very simple, case of two links and three OD pairs the link counts give some information about the relative sizes of the OD pair means. Hence, the result is not completely dependent by the prior chosen. As the network topology grows larger to model more realistic communications network, there are more an more OD pair dependency information in the link counts.

6.2 Further work

Theorem 2 in page 16 is still to be proven in the general case matrix form.

Considering further this kind of static model, where the expected traffic amounts for the OD pairs do not change as a function of time, one way of improving the result is to utilize the covariance matrix in estimation of the mean. To do this we need to assume a relation between mean and variance of the OD pairs. In [2] the authors propose a power law relation where the variance of ith OD pair σ_i^2 would depend linearly on the corresponding mean μ_i raised to power c.

$$\sigma_i^2 \quad = \quad \phi \, \mu_i^c \tag{151}$$

where ϕ, c are parameters to be defined. With this kind of assumption we could get additional information to our estimate.

References

[1] S. Vaton, A. Gravey, "Network tomography: an iterative Bayesian analysis", ITC18, Berlin, August 2003.

[2] J. Cao, D. Davis, S. Vander Wiel, B. Yu, "Time-Varying Network Tomography: Router Link Data" The Journal of American Statistics Association, Vol. 95, No. 452, 2000.