An Introduction To Error Propagation: Derivation, Meaning and Examples of Equation Cy= Fx Cx FxT

Report

Author(s):
Arras, Kai O.

An Introduction To Error Propagation: Derivation, Meaning and Examples of Equation $C_Y = F_X C_X F_X^T$

Kai Oliver Arras

Technical Report

Nº EPFL-ASL-TR-98-01 R3

CONTENTS

1. INTRODUCTION

This report attempts to shed light onto the equation

$$C_Y \;=\; F_X C_X F_X^T, \tag{1}$$

where C_X is a $n \times n$, C_Y a $p \times p$ covariance matrix and F_X some matrix of dimension $p \times n$. In estimation applications like Kalman filtering or probabilistic feature extraction we frequently encounter the pattern $F_X C_X F_X^T$. Many texts in literature introduce this equation without further explanation. But relationship (1), called the *error propagation law*, can be explicitly derived and understood, being important for a comprehension of its underlying approximative character. Here, we want to bridge the gap between these texts and the novice to the world of uncertainty modeling and propagation.

Applications of (1) are e.g. error propagation in model-based vision, Kalman filtering, reliability or probabilistic systems analysis in general.

2. ERROR PROPAGATION: FROM THE BEGINNING

We will first forget the matrix form of (1) and change to a different perspective. Error propagation is the problem of finding the distribution of a function of random variables. Often we have mathematical models of the system of interest (the output as a function of the input and the system components) and we know something about the distribution of the input and the components.
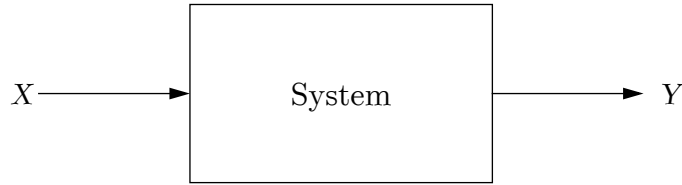

Figure 1: *The simplest case: one input random variable ($N = 1$), and one output random variable ($P = 1$).*

Then, we desire to know the *distribution of the output*, that is the distribution function of Y when $Y = f(X)$ where $f(.)$ is some known function and the distribution function of the random variable X is known[†].

If $f(.)$ is a nonlinear function, the probability distribution function of Y, $p_Y(y)$, becomes quickly very complex, particularly when there is more than one input variable. Although a general method of solution exists (see [BREI70] Chapter 6-6), its complexity can be demonstrated already for simple problems. An approximation of $p_Y(y)$ is therefore desirable. The approximation consists in the propagation of only the first two statistical moments, that is the mean μ_Y and the second (central) moment σ_Y^2, the variance. These moments do not in general describe the distribution of Y. However if Y is assumed to be normally distributed they do[‡].

[†] We do not consider pathological functions $f(.)$ where Y is *not* a random variable – however, they exist
[‡] That is simply one of the favorable properties of the normal distribution.

2.1 A First Expectation

Look at figure 2 where the simple case with one input and one output is illustrated. Suppose that X is normally distributed with mean μ_X and standard deviation σ_X[†]. Now we would like
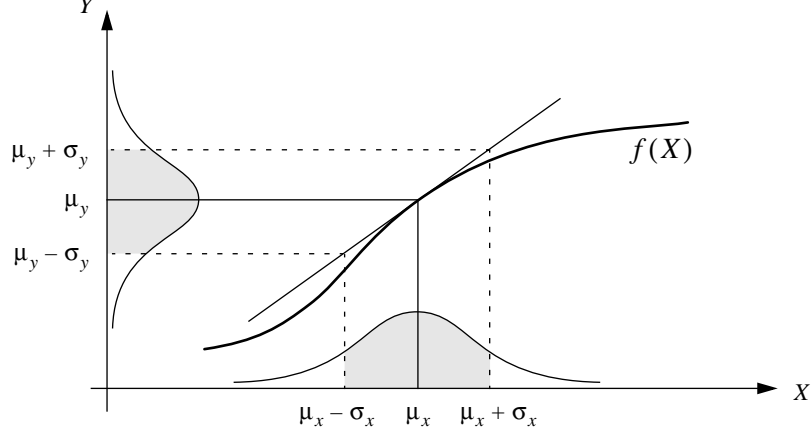

Figure 2: *One-dimensional case of a nonlinear error propagation problem*

to know how the 68% probability interval $[\mu_X - \sigma_X, \mu_X + \sigma_X]$ is propagated through the 'system' $f(.)$.

First of all, from figure 2 it can be seen that if the shaded interval would be mapped onto the y-axis by the original function its shape would be somewhat distorted and the resulting distribution would be asymmetric, certainly not Gaussian anymore. When approximating $f(X)$ by a first-order Taylor series expansion about the point $X = \mu_X$,

$$Y \approx f(\mu_X) + \frac{\partial f}{\partial X}\bigg|_{X = \mu_X} (X - \mu_X), \tag{2}$$

we obtain the linear relationship shown in figure 2 and with that a normal distribution for $p_Y(y)$[‡]. Now we can determine its parameters μ_Y and σ_Y.

$$\mu_Y = f(\mu_X), \tag{3}$$

$$\sigma_Y = \frac{\partial f}{\partial X}\bigg|_{X = \mu_X} \sigma_X. \tag{4}$$

Finally the expectation is rised that in the remainder of this text we will again bump into some generalized form of equation (3) and (4).

At this point, we should not forget that the output distribution, represented by μ_Y and σ_Y, is an approximation of some unknown truth. This truth is impertinently nonlinear, non-normal and asymmetric, thus inhibiting any exact closed form analysis in most cases. We are then supposed to ask the question:

[†] Remember that the standard deviation σ is by definition the distance between the most probable value, μ, and the curve's turning points.

[‡] Another useful property of the normal distribution, worth to be remembered: Gaussian stays Gaussian under linear transformations.

2.2 When is the Approximation a Good One?

Some textbooks write equations (3) and (4) as inequalities. But if the left hand sides denote the parameters of the output distribution which is, by assumption, normal, we can write them as equalities. The first two moments of the true (but unknown) output distribution – let's call them μ_0 and σ_0^2 – are definitely different from these values[†]. Hence

$$\mu_0 \approx \mu_Y \tag{5}$$

$$\sigma_0^2 \approx \sigma_Y^2 \tag{6}$$

It is evident that if $f(.)$ is linear we can write (5) and (6) as equalities. However the following factors affect approximation quality of μ_Y and σ_Y by the actual values y^* and s_y^*.

$$\mu_Y \approx y^* \tag{7}$$

$$\sigma_Y \approx s_y^* \tag{8}$$

Thus they apply for both cases; when $f(.)$ is linear and when it is nonlinear[‡].

- *The guess x^**: In general we do not know the expected value μ_X. In this case, its best (and only available) guess is the actual value of X, for example the current measurement x^*. We then differentiate $f(X)$ at the point $X = x^*$ hoping that x^* is close to $E[X]$ such that y^* is not too far from $E[Y]$.

- *Extent of nonlinearity of $f(.)$*: Equations (3) and (4) are good approximations as long as the first-order Taylor series is a good approximation which is the case if $f(.)$ is not too far from linear within the region that is within one standard deviation of the mean [BREI70]. Some people even ask $f(.)$ to be close to linear within a $\pm 2\sigma$-interval.

The nonlinearity of $f(.)$ and the deviation of x^* from $E[X]$ have a combined effect on σ_Y. If both factors are of high magnitude, the slope at $X = x^*$ can differ strongly and the resulting approximation of σ_Y can be poor. At the other hand, if already one of the them is small, or even both, we can expect s_y^* to be close to σ_Y.

- *The guess s_x^**: There are several possibilities to model the input uncertainty. The model could incorporate some ad hoc assumptions on σ_X or it might rely on an empirically gained relationship to μ_X. Sometimes it is possible to have a physically based model providing 'true' uncertainty information for each realization of X. By systematically following all sources of perturbation during the emergence of x^*, such a model has the desirable property that it accounts for all important factors which influence the outcome x^* and s_x^*.
  In any case, the actual value s_x^* remains a guess of σ_X which is hopefully close to σ_X.

But there is also reason for optimism. The conditions suggested by Figure 2 are exaggerated. Mostly, σ_X *is very small with respect to the range of X*. This makes (2) an approximation of sufficient quality in many practical problems.

[†] Remember that the expected value and the variance (and all other moments) have a general definition, i.e. are independent whether the distribution functions exhibits some nice properties like symmetry. They are also valid for arbitrarily shaped distributions and always quantify the most probable value and its 'spread'.

[‡] It might be possible that in some nonlinear but nice-conditioned cases the approximation effects of non-normality and the other factors compensate each other yielding a very good approximation. Nevertheless, the opposite might also be true.

2.3 The Almost General Case: Approximating the Distribution of $Y = f(X_1, X_2, ..., X_n)$

The next step towards equation (1) is again a practical approximation based on a first-order Taylor series expansion, this time for a multiple-input system. Consider
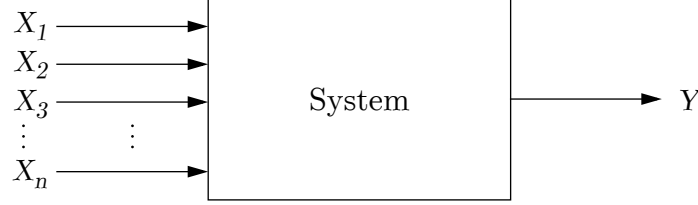

Figure 3: *Error propagation in a multi-input single-output system: $N = n$, $P = 1$.*

$Y = f(X_1, X_2, ..., X_n)$ where the X_i's are n input random variables and Y is represented by its first-order Taylor series expansion about the point $\mu_1, \mu_2, ..., \mu_n$

$$Y \approx f(\mu_1, \mu_2, ..., \mu_n) + \sum_{i=1}^{n} \left[\frac{\partial f}{\partial X_i}(\mu_1, \mu_2, ..., \mu_n) \right][X_i - \mu_i]. \tag{9}$$

Equation (9) is of the form $Y \approx a_0 + \sum a_i(X_i - \mu_i)$ with

$$a_0 = f(\mu_1, \mu_2, ..., \mu_n), \tag{10}$$

$$a_i = \frac{\partial f}{\partial X_i}(\mu_1, \mu_2, ..., \mu_n). \tag{11}$$

As in chapter 2.1, the approximation is linear. The distribution of Y is therefore Gaussian and we have to determine μ_Y and σ_Y^2.

$$\mu_Y = E[Y] = E[a_0 + \sum_i a_i(X_i - \mu_i)] \tag{12}$$

$$= E[a_0] + \sum_i E[a_i X_i] - E[a_i \mu_i] \tag{13}$$

$$= a_0 + \sum_i a_i E[X_i] - a_i E[\mu_i] \tag{14}$$

$$= a_0 + \sum_i a_i \mu_i - a_i \mu_i \tag{15}$$

$$= a_0 \tag{16}$$

$$\mu_Y = f(\mu_1, \mu_2, ..., \mu_n) \tag{17}$$

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] = E[(\sum_i a_i(X_i - \mu_i))^2] \tag{18}$$

$$= E[\sum_i a_i(X_i - \mu_i) \sum_j a_j(X_j - \mu_j)] \tag{19}$$

$$= E[\sum_i a_i^2(X_i - \mu_i)^2 + \sum_i \sum_{i \neq j} a_i a_j(X_i - \mu_i)(X_j - \mu_j)] \tag{20}$$

$$= \sum_i a_i^2 E[(X_i - \mu_i)^2] + \sum_i \sum_{i \neq j} a_i a_j E[(X_i - \mu_i)(X_i - \mu_j)] \tag{21}$$

$$= \sum_i a_i^2 \sigma_i^2 + \sum_i \sum_{i \neq j} a_i a_j \sigma_{ij} \tag{22}$$

$$\sigma_Y^2 = \sum_i \left(\frac{\partial f}{\partial X_i}\right)^2 \sigma_i^2 + \sum_i \sum_{i \neq j} \left(\frac{\partial f}{\partial X_i}\right)\left(\frac{\partial f}{\partial X_j}\right)\sigma_{ij} \tag{23}$$

The vector $\mu_1, \mu_2, ..., \mu_n$ has been omitted. If the X_i's are independent the covariance σ_{ij} disappears, and the resulting approximated variance is

$$\sigma_Y^2 \approx \sum_i \left(\frac{\partial f}{\partial X_i}\right)^2 \sigma_i^2 . \tag{24}$$

This is the moment to validate our expectation from chapter 2.1 for the one-dimensional case. Equation (17) corresponds directly with (3) whereas (23) somewhat contains equation (4).

2.3.1 Addendum to Chapter 2.2

In order to close the discussion on factors affecting approximation quality, we have to consider briefly two aspects which play a role if there is more than one input.

- *Independence of the inputs*: Equation (24) is a good approximation if the stated assumption of independence of all X_i's is valid.

It is finally to be mentioned that, even if the input distributions are not strictly Gaussian, the assumption of the output being normal is often reasonable. This follows from the central limit theorem when the X_i's somehow additively constitute the output Y.

2.4 Getting Really General: Adding $Z = g(X_1, X_2, ..., X_n)$

Often there is not just a single Y which depends on the X_i's but there are more system outputs, that is, more random variables like Y. Suppose our system has an additional output Z with $Z = g(X_1, X_2, ..., X_n)$.
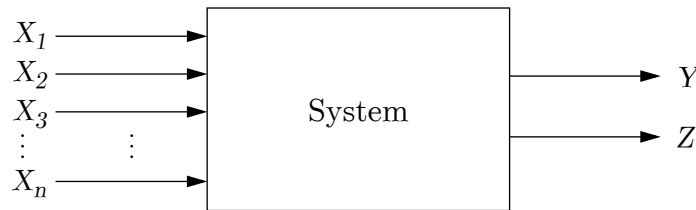
Figure 4: *Error propagation in a multi-input multi-output system: $N = n$, $P = 2$*

Obviously μ_Z and σ_Z^2 can exactly be derived as shown before. The additional aspect which is introduced by Z is the question of the *statistical dependence* of Y and Z which is expressed by their covariance $\sigma_{YZ} = E[(Y - \mu_Y)(Z - \mu_Z)]$. Let's see where we arrive when substituting Y and Z by their first-order Taylor series expansion (9).

$$\sigma_{YZ} = E[(Y - \mu_Y)(Z - \mu_Z)] \tag{25}$$

$$= E[Y \cdot Z] - E[Y]E[Z] \tag{26}$$

$$= E\left[\left(\mu_Y + \sum \frac{\partial f}{\partial X_i}[X_i - \mu_i]\right) \cdot \left(\mu_Z + \sum \frac{\partial g}{\partial X_i}[X_i - \mu_i]\right)\right] - \mu_Y \mu_Z \tag{27}$$

$$= E\left[\mu_Y \mu_Z + \mu_Z \sum \frac{\partial f}{\partial X_i}[X_i - \mu_i] + \mu_Y \sum \frac{\partial g}{\partial X_i}[X_i - \mu_i] + \sum \frac{\partial f}{\partial X_i}[X_i - \mu_i] \sum \frac{\partial g}{\partial X_i}[X_i - \mu_i]\right] - \mu_Y \mu_Z \tag{28}$$

$$= E[\mu_Y \mu_Z] + \mu_Z E\left[\sum \frac{\partial f}{\partial X_i} X_i - \sum \frac{\partial f}{\partial X_i} \mu_i\right] + \mu_Y E\left[\sum \frac{\partial g}{\partial X_i} X_i - \sum \frac{\partial g}{\partial X_i} \mu_i\right]$$
$$+ E\left[\sum \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j}[X_i - \mu_i][X_j - \mu_j]\right] - \mu_Y \mu_Z \tag{29}$$

$$= \mu_Y \mu_Z + \mu_Z \sum \frac{\partial f}{\partial X_i} E[X_i] - \mu_Z \sum \frac{\partial f}{\partial X_i} E[\mu_i] + \mu_Y \sum \frac{\partial g}{\partial X_i} E[X_i] - \mu_Y \sum \frac{\partial g}{\partial X_i} E[\mu_i]$$
$$+ E\left[\sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_i}[X_i - \mu_i]^2 + \sum_{i \neq j} \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j}[X_i - \mu_i][X_j - \mu_j]\right] - \mu_Y \mu_Z \tag{30}$$

$$= \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_i} E[(X_i - \mu_i)^2] + \sum_{i \neq j} \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j} E[(X_i - \mu_i)(X_j - \mu_j)] \tag{31}$$

$$\sigma_{YZ} = \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_i} \sigma_i^2 + \sum_{i \neq j} \sum \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j} \sigma_{ij}^2 \tag{32}$$

If X_i and X_j are independent, the second term, holding their covariance, disappears. Adding more output random variables brings in no new aspects. In the remainder of this text we shall consider $P = 2$ without loss of generality.

3. DERIVATING THE FINAL MATRIX FORM

Now we are ready to return to equation (1). We will now see that we only have to reformulate equations (23) and (32) in order to obtain the initial matrix form. We recall the gradient operator with respect to the n-dimensional vector X

$$\nabla_X = \left[\frac{\partial}{\partial X_1} \frac{\partial}{\partial X_2} \cdots \frac{\partial}{\partial X_n}\right]^T. \tag{33}$$

$f(X)$ is a p-dimensional vector-valued function $f(X) = \left[f_1(X) \ f_1(X) \ \dots \ f_p(X)\right]^T$. The *Jacobian* F_X is defined as the transpose of the gradient of $f(X)$, whereas the gradient is the outer product of ∇_X and $f(X)$

$$F_X = \left[\nabla_X \cdot f(X)^T\right]^T = \begin{bmatrix} f_1(X) \\ f_2(X) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial X_1} \cdots \frac{\partial}{\partial X_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \cdots & \frac{\partial f_1}{\partial X_n} \\ \frac{\partial f_2}{\partial X_1} & \cdots & \frac{\partial f_2}{\partial X_n} \end{bmatrix} \tag{34}$$

F_X has dimension $2 \times n$ in this case, $p \times n$ in general. We introduce the symmetric $n \times n$ *input covariance matrix* C_X which contains all variances and covariances of the input random variables $X_1, X_2, ..., X_n$. If the X_i's are independent all σ_{ij} with $i \neq j$ disappear and C_X is diagonal.

$$C_X = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_n} \\ \sigma_{X_2 X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2 X_n} \\ \vdots & \vdots & & \vdots \\ \sigma_{X_n X_1} & \sigma_{X_n X_2} & \cdots & \sigma_{X_n}^2 \end{bmatrix} \tag{35}$$

We further introduce the symmetric 2×2 *output covariance matrix* C_Y ($p \times p$ in general with outputs $Y_1, Y_2, ..., Y_p$)

$$C_Y = \begin{bmatrix} \sigma_{Y_1}^2 & \sigma_{Y_1 Y_2} \\ \sigma_{Y_2 Y_1} & \sigma_{Y_2}^2 \end{bmatrix} \tag{36}$$

Now we can instantly form equation (1)

$$C_Y = F_X C_X F_X^T \tag{37}$$

$$\begin{bmatrix} \sigma_{Y_1}^2 & \sigma_{Y_1 Y_2} \\ \sigma_{Y_2 Y_1} & \sigma_{Y_2}^2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \cdots & \frac{\partial f_1}{\partial X_n} \\ \frac{\partial f_2}{\partial X_1} & \cdots & \frac{\partial f_2}{\partial X_n} \end{bmatrix} \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_n} \\ \sigma_{X_2 X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2 X_n} \\ \vdots & \vdots & & \vdots \\ \sigma_{X_n X_1} & \sigma_{X_n X_2} & \cdots & \sigma_{X_n}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \frac{\partial f_2}{\partial X_1} \\ \vdots & \vdots \\ \frac{\partial f_1}{\partial X_n} & \frac{\partial f_2}{\partial X_n} \end{bmatrix} \tag{38}$$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial X_1}\sigma_{X_1}^2 + ... + \frac{\partial f_1}{\partial X_n}\sigma_{X_n X_1} & \cdots & \frac{\partial f_1}{\partial X_1}\sigma_{X_1 X_n} + ... + \frac{\partial f_1}{\partial X_n}\sigma_{X_n}^2 \\ \frac{\partial f_2}{\partial X_1}\sigma_{X_1}^2 + ... + \frac{\partial f_2}{\partial X_n}\sigma_{X_n X_1} & \cdots & \frac{\partial f_2}{\partial X_1}\sigma_{X_1 X_n} + ... + \frac{\partial f_2}{\partial X_n}\sigma_{X_n}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \frac{\partial f_2}{\partial X_1} \\ \vdots & \vdots \\ \frac{\partial f_1}{\partial X_n} & \frac{\partial f_2}{\partial X_n} \end{bmatrix} \tag{39}$$

Looks nice. But what has it to do with chapter 2? To answer this question we evaluate the first element $\sigma_{Y_1}^2$, the variance of Y_1:

$$\sigma_{Y_1}^2 = \left(\frac{\partial f_1}{\partial X_1}\right)^2 \sigma_{X_1}^2 + \frac{\partial f_1}{\partial X_2}\frac{\partial f_1}{\partial X_1}\sigma_{X_2 X_1} + ... + \frac{\partial f_1}{\partial X_n}\frac{\partial f_1}{\partial X_1}\sigma_{X_n X_1} + \frac{\partial f_1}{\partial X_1}\frac{\partial f_1}{\partial X_2}\sigma_{X_1 X_2} + \left(\frac{\partial f_1}{\partial X_2}\right)^2 \sigma_{X_2}^2 + ... + \frac{\partial f_1}{\partial X_n}\frac{\partial f_1}{\partial X_2}\sigma_{X_n X_2}$$

$$+ ... + \frac{\partial f_1}{\partial X_1}\frac{\partial f_1}{\partial X_n}\sigma_{X_1 X_n} + \frac{\partial f_1}{\partial X_2}\frac{\partial f_1}{\partial X_n}\sigma_{X_2 X_n} + ... + \left(\frac{\partial f_1}{\partial X_n}\right)^2 \sigma_{X_n}^2 \tag{40}$$

$$= \sum_i \left(\frac{\partial f_1}{\partial X_i}\right)^2 \sigma_{X_i}^2 + \sum_{i \neq j} \sum \frac{\partial f_1}{\partial X_i}\frac{\partial f_1}{\partial X_j}\sigma_{X_i X_j} \tag{41}$$

If we now reintroduce the notation of chapter 2.3, that is, $f_1(X) = f(X)$, $Y_1 = Y$, and $\sigma_{X_i} = \sigma_i$, we see that (41) equals exactly equation (23). Assuming the reader being a notorious skeptic, we will also look at the off-diagonal element $\sigma_{Y_1 Y_2}$, the covariance of Y and Z:

$$\sigma_{Y_1 Y_2} = \frac{\partial f_1}{\partial X_1}\frac{\partial f_2}{\partial X_1}\sigma_{X_1}^2 + \frac{\partial f_1}{\partial X_2}\frac{\partial f_2}{\partial X_1}\sigma_{X_2 X_1} + ... + \frac{\partial f_1}{\partial X_n}\frac{\partial f_2}{\partial X_1}\sigma_{X_n X_1} + \frac{\partial f_1}{\partial X_1}\frac{\partial f_2}{\partial X_2}\sigma_{X_1 X_2} + \frac{\partial f_1}{\partial X_2}\frac{\partial f_2}{\partial X_2}\sigma_{X_2}^2 + ... + \frac{\partial f_1}{\partial X_n}\frac{\partial f_2}{\partial X_2}\sigma_{X_n X_2}$$

$$+ ... + \frac{\partial f_1}{\partial X_1}\frac{\partial f_2}{\partial X_n}\sigma_{X_1 X_n} + \frac{\partial f_1}{\partial X_2}\frac{\partial f_2}{\partial X_n}\sigma_{X_2 X_n} + ... + \frac{\partial f_1}{\partial X_n}\frac{\partial f_2}{\partial X_n}\sigma_{X_n}^2 \tag{42}$$

$$= \sum_i \frac{\partial f_1}{\partial X_i}\frac{\partial f_2}{\partial X_i}\sigma_{X_i}^2 + \sum_{i \neq j}\sum \frac{\partial f_1}{\partial X_i}\frac{\partial f_2}{\partial X_j}\sigma_{X_i X_j} \qquad (43)$$

Again, by substituting $f_1(X)$ by $f(X)$, $f_2(X)$ by $g(X)$ and σ_{X_i} by σ_i, equation (43) correspond exactly to the previously derived equation (32) for σ_{YZ}.

We were obviously able, having started from a simple one-dimensional error propagation problem, to derive the error propagation law $C_Y = F_X C_X F_X^T$. Putting the results together yielded its widely used matrix form (1).

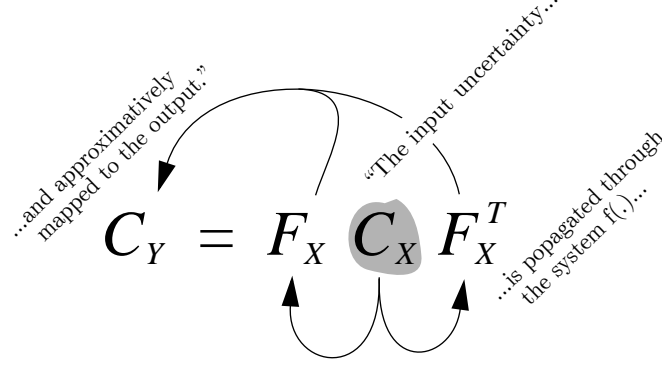Now we can also understand the informal interpretation of figure 5.

Figure 5: *Interpretation of the error propagation law in its matrix form*

4. EXAMPLES

4.1 Probabilistic Line Extraction From Noisy 1D Range Data

Model-based vision where geometric primitives are the sought image features is a good example for uncertainty propagation. Suppose the segmentation problem has already been solved, that is, the set of inlier points with respect to the model is known. Suppose further that the regression equations for the model fit to the points have a closed-form solution – which is the case when fitting straight lines. And suppose finally, that the measurement uncertainties of the data points are known as well. Then we can directly apply the error propagation law in order to get the output uncertainties of the line parameters.
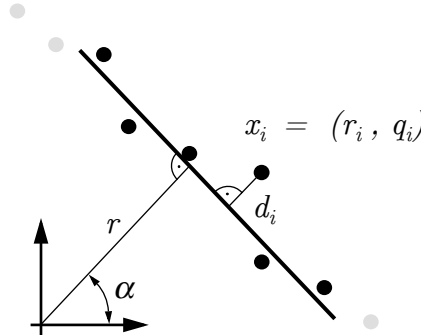

Figure 6: *Estimating a line in the least squares sense. The model parameters r (length of the perpendicular) and α (its angle to the abscissa) describe uniquely a line.*

Suppose n measurement points in polar coordinates $x_i = (\rho_i, \theta_i)$ are given and modeled as random variables $X_i = (P_i, Q_i)$ with $i = 1, ..., n$. Each point is independently affected by Gaussian noise in both coordinates.

$$P_i \sim N(\rho_i, \sigma_{\rho_i}^2) \tag{44}$$

$$Q_i \sim N(\theta_i, \sigma_{\theta_i}^2) \tag{45}$$

$$E[P_i \cdot P_j] = E[P_i]E[P_j] \qquad \forall \, i, j = 1, ..., n \tag{46}$$

$$E[Q_i \cdot Q_j] = E[Q_i]E[Q_j] \qquad \forall \, i, j = 1, ..., n \tag{47}$$

$$E[P_i \cdot Q_j] = E[P_i]E[Q_j] \qquad \forall \, i, j = 1, ..., n \tag{48}$$

Now we want to find the line $x\cos\alpha + y\sin\alpha - r = 0$ where $x = \rho\cos(\theta)$ and $y = \rho\sin(\theta)$ yielding $\rho\cos\theta\cos\alpha + \rho\sin\theta\sin\alpha - r = 0$ and with that, the line model

$$\rho\cos(\theta - \alpha) - r = 0. \tag{49}$$

This model minimizes the *orthogonal* distances from the points to the line. It is important to note that fitting models to data in *some* least square sense yields not a satisfying geometric solution in general. It is crucial to know *which* error is minimized by the fit equations. A good illustration is the paper of [GAND94] where several algorithms for fitting circles and ellipses are presented which minimize algebraic and geometric distances. The geometric variety of solutions for the same set of points demonstrate the importance of this knowledge if geometric meaningful results are required.

The orthogonal distance d_i of a point (ρ_i, θ_i) to the line is just

$$\rho_i\cos(\theta_i - \alpha) - r = d_i. \tag{50}$$

Let S be the (unweighted) sum of squared errors.

$$S = \sum_i d_i^2 = \sum_i (\rho_i\cos(\theta_i - \alpha) - r)^2 \tag{51}$$

The model parameters (α, r) are now found by solving the nonlinear equation system

$$\frac{\partial S}{\partial \alpha} = 0 \qquad \frac{\partial S}{\partial r} = 0. \tag{52}$$

Suppose further that for each point a variance σ_i^2 modelling the uncertainty in radial and angular direction is given a priori or can be measured. This variance will be used to determine a weight w_i for each single point, e.g.

$$w_i = 1/\sigma_i^2 \, ^\dagger. \tag{53}$$

Then, equation (51) becomes

$$S = \sum w_i d_i^2 = \sum w_i(\rho_i\cos(\theta_i - \alpha) - r)^2. \tag{54}$$

---

† The issue of determining an adequate weight when σ_i (and perhaps some additional information) is given is complex in general and beyond the scope of this text. See [CARR88] for a careful treatment.

It can be shown (see Appendix A) that the solution of (52) in the *weighted* least square sense[†] is

$$\alpha = \frac{1}{2}\text{atan2}\left(\frac{\sum w_i\rho_i^2\sin 2\theta_i - \frac{2}{\sum w_i}\sum\sum w_iw_j\rho_i\rho_j\cos\theta_i\sin\theta_j}{\sum w_i\rho_i^2\cos 2\theta_i - \frac{1}{\sum w_i}\sum\sum w_iw_j\rho_i\rho_j\cos(\theta_i + \theta_j)}\right) \tag{55}$$

$$r = \frac{\sum w_i\rho_i\cos(\theta_i - \alpha)}{\sum w_i}. \tag{56}$$

Now we would like to know how the uncertainties of the measurements propagate through 'the system' (55), (56). See figure 7.
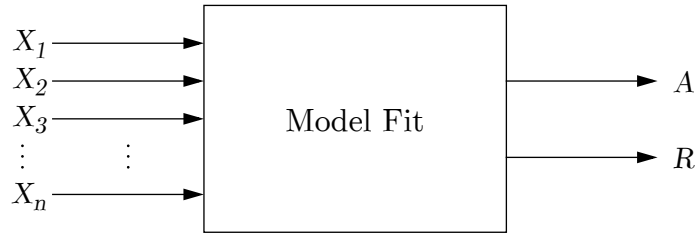

Figure 7: *When extracting lines from noisy measurement points X_i, the model fit module produces line parameter estimates, modeled as random variables A, R. It is then interesting to know the variability of these parameters as a function of the noise at the input side.*

This is where we simply apply equation (1). We are looking for the 2×2 output covariance matrix

$$C_{AR} = \begin{bmatrix} \sigma_A^2 & \sigma_{AR} \\ \sigma_{AR} & \sigma_R^2 \end{bmatrix}, \tag{57}$$

given the $2n \times 2n$ input covariance matrix

$$C_X = \begin{bmatrix} C_P & 0 \\ 0 & C_Q \end{bmatrix} = \begin{bmatrix} diag(\sigma_{\rho_i}^2) & 0 \\ 0 & diag(\sigma_{\theta_i}^2) \end{bmatrix} \tag{58}$$

and the system relationships (55) and (56). Then by calculating the Jacobian

$$F_{PQ} = \begin{bmatrix} \dfrac{\partial\alpha}{\partial P_1} & \dfrac{\partial\alpha}{\partial P_2} & \cdots & \dfrac{\partial\alpha}{\partial P_n} & \dfrac{\partial\alpha}{\partial Q_1} & \dfrac{\partial\alpha}{\partial Q_2} & \cdots & \dfrac{\partial\alpha}{\partial Q_n} \\ \dfrac{\partial r}{\partial P_1} & \dfrac{\partial r}{\partial P_2} & \cdots & \dfrac{\partial r}{\partial P_n} & \dfrac{\partial r}{\partial Q_1} & \dfrac{\partial r}{\partial Q_2} & \cdots & \dfrac{\partial r}{\partial Q_n} \end{bmatrix} \tag{59}$$

[†] We follow here the notation of [DRAP88] and distinguish a weighted least squares problem if C_X is diagonal (input errors are mutually independent) and a generalized least squares problem if C_X is non-diagonal.

we can instantly form the error propagation equation (60) yielding the sought C_{AR}.

$$C_{AR} = F_{PQ} C_X F_{PQ}^T \tag{60}$$

Appendix A is concerned about the step-by-step derivation of the fit equations (55) and (56), whereas in Appendix B equation (60) is once more derived. Under the assumption of negligible angular uncertainties, implementable expressions for the elements of C_{AR} are determined, also in a step-by-step manner.

4.2 Kalman Filter Based Mobile Robot Localization: The Measurement Model

The measurement model in a Kalman filter estimation problem is another place where we encounter the error propagation law. The reader is assumed to be more or less familiar with the context of mobile robot localization, Kalman filtering and the notation used in [BAR93] or [LEON92].

In order to reduce the unbounded growth of odometry errors the robot is supposed to update its pose by some sort of external referencing. This is achieved by matching predicted environment features with observed ones and estimating the vehicle pose in some sense with the set of matched features. The prediction is provided by an a priori map which contains the position of all environment features in global map coordinates. In order to establish correct correspondence of observed and stored features, the robot predicts the position of all currently visible features in the sensor frame. This is done by the measurement model

$$\hat{z}(k+1|k) = h(\hat{x}(k+1|k)) + w(k+1). \tag{61}$$

The measurement model gets the predicted robot pose $\hat{x}(k+1|k)$ as input and 'asks' the map which features are visible at the current location and where they are supposed to appear when starting the observation. The map evaluates all visible features and returns their transformed positions in the vector $\hat{z}(k+1|k)$. The measurement model is therefore a world-to-robot-to-sensor frame transformation[†].

However, due to nonsystematic odometry errors, the robot position is uncertain and due to imperfect sensors, the observation is uncertain. The former is represented by the (predicted) state covariance matrix $P(k+1|k)$ and the latter by the sensor noise model $w(k+1) \sim N(0, R(k+1))$. They are assumed to be independent. Sensing uncertainty $R(k+1)$ affects the observation *directly*, whereas vehicle position uncertainty $P(k+1|k)$ – which is given in world map coordinates – will *propagate* through the frame transformations world-to-robot-to-sensor $h(.)$, linearized about the prediction $\hat{x}(k+1|k)$. Then the observation is made and the matching of predicted and observed features can be performed in the sensor frame yielding the set of matched features. The remaining position uncertainty of the matched features given all observations up to and including time k,

$$S(k+1) = \text{cov}[z(k+1)|Z^k], \tag{62}$$

[†] Note that $h(.)$ is assumed to be 'intelligent', that is, it contains both, the mapping $\hat{x}(k+1|k) \to m_j$ (with m_j as the position vector of feature number j) and the world-to-sensor frame transformation of all visible m_j for the current prediction $\hat{x}(k+1|k)$. This is in contrast to e.g. [LEON92], where solely the frame transformation is done by $h(\hat{x}(k+1|k), m_j)$ and the mapping $\hat{x}(k+1|k) \to m_j$ is somewhere else.

is the superimposed uncertainty of observation and the propagated one from the robot pose,

$$S(k+1) = \nabla h P(k+1|k) \nabla h + R(k+1). \tag{63}$$

$S(k+1)$ is also called *measurement prediction covariance* or *innovation covariance.*

5. EXERCISES

As stated in the introduction, the report has educational purposes and accompanies a lecture on autonomous systems in the microengineering departement at EPFL. Thus, the audience of this report are people not yet too familarized with the field of uncertainty treatment. Some propositions for exercises are given:

- Let them do the derivation of μ_Y (equations (12) to (17)) and σ_Y^2 (equations (18) to (23)) given a few rules for the expected value.

- Let them do the derivation of σ_{YZ} (equations (25) to (32)) or, in the context of example 1, equations (93) to (101) of Appendix B for σ_{AR}. Some rules for the expected value and double sums might be helpful.

- Let them make an illustration of each factor affecting approximation quality discussed in chapter 2.2 with drawings like figure 2.

- If least squares estimation in a more general sense is the issue, derivating the fit equations for a regression problem is quite instructive. The standard case, linear in the model parameters, and with uncertainties in only one variable is much simpler than the derivation of example 1 in Appendix A. Additionally the output covariance matrix can be determined with (1).

LITERATURE

[BREI70]   A.M. Breipohl, *Probabilistic Systems Analysis: An Introduction to Probabilistic Models, Decisions, and Applications of Random Processes*, John Wiley & Sons, 1970.

[GAND94]   W. Gander, G. H. Golub, R. Strebel, "Least-Squares Fitting of Circles and Ellipses", BIT, vol. 34, no. 4, p. 558-78, Dec. 1994.

[DRAP88]   N.R. Draper, H. Smith, *Applied Regression Analysis*, 3rd edition, John Wiley & Sons, 1988.

[CARR88]   R. J. Carroll, D. Ruppert, *Transformation and Weighting in Regression*, Chapman and Hall, 1988.

[BAR93]   Y. Bar-Shalom, X.-R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, 1993.

[LEON92]   J.J. Leonard, H.F. Durrant-Whyte, *Directed Sonar Sensing for Mobile Robot Navigation*, Kluwer Academic Publishers, 1992.

[ARRAS97]   K.O. Arras, R.Y. Siegwart, "Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building", Proceedings of SPIE, Mobile Robotics XII, Vol. 3210, p. 42-53, 1997.

Appendix A: Finding the Line Parameters r and α in the Weighted Least Square Sense

Consider the nonlinear equation system

$$\frac{\partial S}{\partial r} = 0 \tag{64}$$

$$\frac{\partial S}{\partial \alpha} = 0 \tag{65}$$

where S is the weighted sum of squared errors

$$S = \sum_{i=1}^{n} w_i(\rho_i \cos\theta_i \cos\alpha + \rho_i \sin\theta_i \sin\alpha - r)^2. \tag{66}$$

We start solving the system (64), (65) by working out parameter r.

$$\frac{\partial S}{\partial r} = 0 = \sum 2w_i(\rho_i \cos\theta_i \cos\alpha + \rho_i \sin\theta_i \sin\alpha - r)(-1) \tag{67}$$

$$= -2\sum w_i\rho_i(\cos\theta_i \cos\alpha + \sin\theta_i \sin\alpha) + 2\sum w_i r \tag{68}$$

$$= -2\sum w_i\rho_i\cos(\theta_i - \alpha) + 2r\sum w_i \tag{69}$$

$$2r\sum w_i = 2\sum w_i\rho_i\cos(\theta_i - \alpha) \tag{70}$$

$$r = \frac{\sum w_i\rho_i\cos(\theta_i - \alpha)}{\sum w_i} \tag{71}$$

Parameter α is slightly more complicated. We introduce the following notation

$$\cos\theta_i = c_i \qquad \sin\theta_i = s_i. \tag{72}$$

$$\frac{\partial S}{\partial \alpha} = \sum 2w_i(\rho_i c_i \cos\alpha + \rho_i s_i \sin\alpha - r)\frac{\partial}{\partial \alpha}(\rho_i c_i \cos\alpha + \rho_i s_i \sin\alpha - r) \tag{73}$$

$$= 2\sum w_i\left(\rho_i c_i \cos\alpha + \rho_i s_i \sin\alpha - \frac{1}{\sum w_j}\sum w_j\rho_j\cos(\theta_j - \alpha)\right)\left(-\rho_i c_i \sin\alpha + \rho_i s_i \cos\alpha - \frac{\partial}{\partial \alpha}r\right) \tag{74}$$

$$= 2\sum w_i\left(\rho_i c_i \cos\alpha + \rho_i s_i \sin\alpha - \frac{1}{\sum w_j}\sum w_j\rho_j\cos(\theta_j - \alpha)\right)\left(\rho_i s_i \cos\alpha - \rho_i c_i \sin\alpha - \frac{1}{\sum w_j}\sum w_j\rho_j\sin(\theta_j - \alpha)\right) \tag{75}$$

$$= 2\sum w_i\Big[\rho_i^2 c_i s_i \cos^2\alpha - \rho_i^2 c_i^2 \cos\alpha\sin\alpha - \frac{1}{\sum w_j}\rho_i c_i \cos\alpha\sum w_j\rho_j\sin(\theta_j - \alpha) + \rho_i^2 s_i^2 \cos\alpha\sin\alpha$$

$$- \rho_i^2 s_i c_i \sin^2\alpha - \frac{1}{\sum w_j}\rho_i s_i \sin\alpha\sum w_j\rho_j\sin(\theta_j - \alpha) - \frac{1}{\sum w_j}\sum w_j\rho_j\cos(\theta_j - \alpha)\rho_i s_i \cos\alpha$$

$$+ \frac{1}{\sum w_j}\sum w_j\rho_j\cos(\theta_j - \alpha)\rho_i s_i \cos\alpha + \frac{1}{(\sum w_j)^2}\sum w_j\rho_j\cos(\theta_j - \alpha)\sum w_j\rho_j\sin(\theta_j - \alpha)\Big] \tag{76}$$

$$= 2\sum w_i\Big[\rho_i^2 c_i s_i(\cos^2\alpha - \sin^2\alpha) - \rho_i^2\cos\alpha\sin\alpha(c_i^2 - s_i^2) - \frac{1}{\sum w_j}\sum w_j\rho_i\rho_j c_i \cos\alpha\sin(\theta_j - \alpha)$$

$$- \frac{1}{\sum w_j}\sum w_j\rho_i\rho_j s_i \sin\alpha\sin(\theta_j - \alpha) - \frac{1}{\sum w_j}\sum w_j\rho_i\rho_j s_i \cos\alpha\cos(\theta_j - \alpha)$$

$$+ \frac{1}{\sum w_j}\sum w_j\rho_i\rho_j c_i \sin\alpha\cos(\theta_j - \alpha) + \frac{1}{(\sum w_j)^2}\sum_j\sum_k w_j w_k\rho_j\rho_k\cos(\theta_j - \alpha)\sin(\theta_k - \alpha)\Big] \tag{77}$$

$$= 2\cos 2\alpha \sum w_i\rho_i^2 c_i s_i - 2\cos\alpha\sin\alpha \sum w_i\rho_i^2 c_{2i} + \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j(c_i\cos\alpha + s_i\sin\alpha)(s_j\cos\alpha - c_j\sin\alpha)$$

$$+ \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j[-c_i\cos\alpha(s_j\cos\alpha - c_j\sin\alpha) - s_i\sin\alpha(s_j\cos\alpha - c_j\sin\alpha)$$

$$+ s_i\cos\alpha(c_j\cos\alpha + s_j\sin\alpha) + c_i\sin\alpha(c_j\cos\alpha + s_j\sin\alpha)] \tag{78}$$

$$= \cos 2\alpha \sum w_i\rho_i^2 s_{2i} - \sin 2\alpha \sum w_i\rho_i^2 c_{2i}$$

$$+ \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j[c_i s_j\cos^2\alpha - c_i c_j\sin\alpha\cos\alpha + s_i s_j\sin\alpha\cos\alpha - s_i c_j\sin^2\alpha - c_i s_j\cos^2\alpha$$

$$+ c_i c_j\sin\alpha\cos\alpha - s_i s_j\sin\alpha\cos\alpha + s_i c_j\sin^2\alpha - s_i c_j\cos^2\alpha$$

$$- s_i s_j\sin\alpha\cos\alpha + c_i c_j\sin\alpha\cos\alpha + c_i s_j\sin^2\alpha] \tag{79}$$

$$= \cos 2\alpha \sum w_i\rho_i^2 s_{2i} - \sin 2\alpha \sum w_i\rho_i^2 c_{2i}$$

$$+ \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j(-s_i c_j\cos^2\alpha - s_i s_j\sin\alpha\cos\alpha + c_i c_j\sin\alpha\cos\alpha + c_i s_j\sin^2\alpha) \tag{80}$$

$$= \cos 2\alpha \sum w_i\rho_i^2 s_{2i} - \sin 2\alpha \sum w_i\rho_i^2 c_{2i}$$

$$+ \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j[\sin\alpha\cos\alpha(c_i c_j - s_i s_j) + c_i s_j\sin^2\alpha - s_i c_j\cos^2\alpha] \tag{81}$$

$$= \cos 2\alpha \sum w_i\rho_i^2 s_{2i} - \sin 2\alpha \sum w_i\rho_i^2 c_{2i} + \frac{2}{\sum w_i}\sin\alpha\cos\alpha\sum\sum w_i w_j\rho_i\rho_j c_{i+j}$$

$$+ \frac{2}{\sum w_i}\sin^2\alpha\sum\sum w_i w_j\rho_i\rho_j c_i s_j - \frac{2}{\sum w_i}\cos^2\alpha\sum\sum w_i w_j\rho_i\rho_j s_i c_j \tag{82}$$

$$= \cos 2\alpha \sum w_i\rho_i^2 s_{2i} - \sin 2\alpha \sum w_i\rho_i^2 c_{2i} + \frac{1}{\sum w_i}\sin 2\alpha\sum\sum w_i w_j\rho_i\rho_j c_{i+j}$$

$$- \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j c_i s_j(\cos^2\alpha - \sin^2\alpha) \tag{83}$$

$$= \cos 2\alpha\left(\sum w_i\rho_i^2 s_{2i} - \frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j c_i s_j\right) - \sin 2\alpha\left(\sum w_i\rho_i^2 c_{2i} - \frac{1}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j c_{i+j}\right) \tag{84}$$

From (84) we can obtain the result for α or $\tan 2\alpha$ respectively.

$$\frac{\sin 2\alpha}{\cos 2\alpha} = \frac{\frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j c_i s_j - \sum w_i\rho_i^2 s_{2i}}{\frac{1}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j c_{i+j} - \sum w_i\rho_i^2 c_{2i}} \tag{85}$$

$$\tan 2\alpha = \frac{\frac{2}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j\cos\theta_i\sin\theta_j - \sum w_i\rho_i^2\sin 2\theta_i}{\frac{1}{\sum w_i}\sum\sum w_i w_j\rho_i\rho_j\cos(\theta_i + \theta_j) - \sum w_i\rho_i^2\cos 2\theta_i} \tag{86}$$

Equation (86) contains double sums which may not be fully evaluated. Due to the symmetry of trigonometric functions the corresponding off-diagonal elements can be added and thus simplifies calculation. For the final result (87), the four-quadrant arc tangent has been taken. This solution, (71) and (87), generates sometimes (α, r)-pairs with negative r values. They must be detected in order to change the sign of r and to add π to the corresponding α. All α-values lie then in the interval $-\pi/2 < \alpha \le 3\pi/2$.

$$\alpha = \frac{1}{2}\text{atan2}\left(\frac{\frac{2}{\sum w_i}\sum_{i<j}\sum w_i w_j\rho_i\rho_j\sin(\theta_i + \theta_j) + \frac{1}{\sum w_i}\sum(w_i - \Sigma w_j)w_i\rho_i^2\sin 2\theta_i}{\frac{2}{\sum w_i}\sum_{i<j}\sum w_i w_j\rho_i\rho_j\cos(\theta_i + \theta_j) + \frac{1}{\sum w_i}\sum(w_i - \Sigma w_j)w_i\rho_i^2\cos 2\theta_i}\right) \tag{87}$$

Appendix B: Deriving the Covariance Matrix for r and α

The model parameter α and r are just half the battle. Besides these estimates of the mean position of the line we would also like to have a measure for its uncertainty. According to the error propagation law (1), an approximation of the output uncertainty, represented by C_{AR}, is subsequently determined.

At the input side, mutually independent uncertainties in radial direction only are assumed. The $2n \times 1$ input random vector $X^T = [P^T \ Q^T]$ consists of the random vector $P = [P_1, P_2, ..., P_n]^T$ denoting the variables of the measured radii, and the random vector $Q = [Q_1, Q_2, ..., Q_n]^T$ holding the corresponding angular variates. The $2n \times 2n$ input covariance matrix C_X is therefore of the form

$$C_X = \begin{bmatrix} C_P & \mathbf{0} \\ \mathbf{0} & C_Q \end{bmatrix} = \begin{bmatrix} diag(\sigma_{\rho_i}^2) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{88}$$

We represent both output random variables A, R by their first-order Taylor series expansion about the mean $\mu^T = [\mu_\rho^T \ \mu_\theta^T]$. The vector μ has dimension $2n \times 1$ and is composed of the two $n \times 1$ mean vectors $\mu_\rho = [\mu_{\rho_1}, \mu_{\rho_2}, ..., \mu_{\rho_n}]^T$ and $\mu_\theta = [\mu_{\theta_1}, \mu_{\theta_2}, ..., \mu_{\theta_n}]^T$.

$$A \approx \alpha(\mu^*) + \sum \frac{\partial \alpha}{\partial P_i}\bigg|_{X=\mu^*} [P_i - \mu_{\rho_i}^*] + \sum \frac{\partial \alpha}{\partial Q_i}\bigg|_{X=\mu^*} [Q_i - \mu_{\theta_i}^*] \tag{89}$$

$$R \approx r(\mu^*) + \sum \frac{\partial r}{\partial P_i}\bigg|_{X=\mu^*} [P_i - \mu_{\rho_i}^*] + \sum \frac{\partial r}{\partial Q_i}\bigg|_{X=\mu^*} [Q_i - \mu_{\theta_i}^*] \tag{90}$$

The relationships $r(.)$ and $\alpha(.)$ correspond to the results of Appendix A, equations (71) and (87). Referring to the discussion of chapter 2.2, we do not know μ in advance and its best available guess is the actual value of the measurements vector μ^*.

It has been shown that under the assumption of independence of P and Q, the following holds

$$\sigma_A^2 \approx \sum \left(\frac{\partial \alpha}{\partial P_i}\right)^2 \sigma_{\rho_i}^2 + \sum \left(\frac{\partial \alpha}{\partial Q_i}\right)^2 \sigma_{\theta_i}^2 \tag{91}$$

$$\sigma_R^2 \approx \sum \left(\frac{\partial r}{\partial P_i}\right)^2 \sigma_{\rho_i}^2 + \sum \left(\frac{\partial r}{\partial Q_i}\right)^2 \sigma_{\theta_i}^2 \tag{92}$$

We further want to know the covariance σ_{AR} under the abovementioned assumption of negligible angular uncertainties:

$$COV[A, R] = E[A \cdot R] - E[A]E[R] \tag{93}$$

$$= E\left[\left(\alpha + \sum \frac{\partial \alpha}{\partial P_i}[P_i - \mu_{\rho_i}]\right) \cdot \left(r + \sum \frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}]\right)\right] - \alpha r \tag{94}$$

$$= E\left[\alpha r + r\sum \frac{\partial \alpha}{\partial P_i}[P_i - \mu_{\rho_i}] + \alpha\sum \frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}] + \sum \frac{\partial \alpha}{\partial P_i}[P_i - \mu_{\rho_i}]\sum \frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}]\right] - \alpha r \tag{95}$$

$$= E[\alpha r] + rE\left[\sum \frac{\partial \alpha}{\partial P_i}P_i - \sum \frac{\partial \alpha}{\partial P_i}\mu_{\rho_i}\right] + \alpha E\left[\sum \frac{\partial r}{\partial P_i}P_i - \sum \frac{\partial r}{\partial P_i}\mu_{\rho_i}\right]$$

$$+ E\left[\sum\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}][P_j - \mu_{\rho_j}]\right] - \alpha r \tag{96}$$

$$= \alpha r + r\sum\frac{\partial}{\partial P_i}(\alpha)E[P_i] - r\sum\frac{\partial\alpha}{\partial P_i}E[\mu_{\rho_i}] + \alpha\sum\frac{\partial r}{\partial P_i}E[P_i] - \alpha\sum\frac{\partial}{\partial P_i}(r)E[\mu_{\rho_i}]$$

$$+ E\left[\sum_{i\neq j}\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}][P_j - \mu_{\rho_j}] + \sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}]^2\right] - \alpha r \tag{97}$$

$$= E\left[\sum_{i\neq j}\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}(P_iP_j - P_i\mu_{P_j} - P_j\mu_{P_i} + \mu_{P_i}\mu_{P_j})\right] + E\left[\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}[P_i - \mu_{\rho_i}]^2\right] \tag{98}$$

$$= \sum_{i\neq j}\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}E[P_iP_j - P_i\mu_{P_j} - P_j\mu_{P_i} + \mu_{P_i}\mu_{P_j}] + \sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}E[(P_i - \mu_{\rho_i})^2] \tag{99}$$

$$= \sum_{i\neq j}\sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}E[P_iP_j - P_i\mu_{P_j} - P_j\mu_{P_i} + \mu_{P_i}\mu_{P_j}] + \sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}\sigma^2_{\rho_i} \tag{100}$$

Since P_i and P_j are independent, the expected value of the bracketed expression disappears. Hence

$$\sigma_{AR} = \sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}\sigma^2_{\rho_i} \tag{101}$$

If, however, the input uncertainty model provides non-negligible angular variances, it is easy to show that under the independency assumption of P and Q the expression keeping track of Q can be simply added to yield

$$\sigma_{AR} = \sum\frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}\sigma^2_{\rho_i} + \sum\frac{\partial\alpha}{\partial Q_i}\frac{\partial r}{\partial Q_i}\sigma^2_{\theta_i}. \tag{102}$$

As demonstrated in chapter 3, the results (91), (92) and (102) can also be obtained in the more compact but less intuitive form of equation (1). Let

$$F_{PQ} = \begin{bmatrix} F_P & F_Q \end{bmatrix} = \begin{bmatrix} \frac{\partial\alpha}{\partial P} & \frac{\partial\alpha}{\partial Q} \\ \frac{\partial r}{\partial P} & \frac{\partial r}{\partial Q} \end{bmatrix} = \begin{bmatrix} \frac{\partial\alpha}{\partial P_1} & \frac{\partial\alpha}{\partial P_2} & \cdots & \frac{\partial\alpha}{\partial P_n} & \frac{\partial\alpha}{\partial Q_1} & \frac{\partial\alpha}{\partial Q_2} & \cdots & \frac{\partial\alpha}{\partial Q_n} \\ \frac{\partial r}{\partial P_1} & \frac{\partial r}{\partial P_2} & \cdots & \frac{\partial r}{\partial P_n} & \frac{\partial r}{\partial Q_1} & \frac{\partial r}{\partial Q_2} & \cdots & \frac{\partial r}{\partial Q_n} \end{bmatrix} \tag{103}$$

the composed $p \times 2n$ Jacobian matrix containing all partial derivates of the model parameters with respect to the input random variables about the guess $\mu*$. Then, the sought covariance matrix C_{AR} can be rewritten as

$$C_{AR} = F_{PQ}C_XF_{PQ}^T. \tag{104}$$

Under the conditions which lead to equation (102), the right hand side can be decomposed yielding

$$C_{AR} = F_PC_PF_P^T + F_QC_QF_Q^T \tag{105}$$

B.1 Practical Considerations

We determine implementable expressions for the elements of covariance matrix C_{AR}. Under the assumption of negligible angular uncertainties, concrete expressions of σ_A^2, σ_R^2 and σ_{AR} will be derived. We must furthermore keep in mind that our problem might be a real time problem, requiring an efficient implementation. Expression are therefore sought which minimize the numer of floating point operations, e.g. by reuse of already computed subexpressions. Although calculating with weights does not add much difficulty, we will omit them in this chapter[†].

For the sake of brevity we introduce the following notation

$$\alpha = \frac{1}{2}\operatorname{atan2}\left(\frac{D}{N}\right) \tag{106}$$

with

$$N = \frac{2}{n}\sum\sum P_i P_j \cos Q_i \sin Q_j - \sum P_i^2 \sin 2Q_i, \tag{107}$$

$$D = \frac{1}{n}\sum\sum P_i P_j \cos(Q_i + Q_j) - \sum P_i^2 \cos 2Q_i. \tag{108}$$

We will use the result that the parameters α (equation (87)) and r (equation (71)) can also be written in Cartesian form, where $x_i = \rho_i \cos\theta_i$ and $y_i = \rho_i \sin\theta_i$:

$$\alpha = \frac{1}{2}\operatorname{atan2}\left\{\frac{-2\sum(\bar{y} - y_i)(\bar{x} - x_i)}{\sum(\bar{y} - y_i)^2 - (\bar{x} - x_i)^2}\right\}, \tag{109}$$

$$r = \bar{x}\cos\alpha + \bar{y}\sin\alpha. \tag{110}$$

They use the means

$$\bar{x} = 1/n \cdot \sum x_i \qquad \bar{y} = 1/n \cdot \sum y_i. \tag{111}$$

From equations (91), (92) and (101) we see that the covariance matrix is defined when both partial derivates of the parameters with respect to P_i are given. Let us start with the derivate of α.

$$\frac{\partial\alpha}{\partial P_i} = \frac{1}{2} \cdot \frac{1}{1 + N^2/D^2} \cdot \frac{\frac{\partial D}{\partial P_i}N - \frac{\partial N}{\partial P_i}D}{D^2} = \frac{1}{2} \cdot \frac{\frac{\partial D}{\partial P_i}N - \frac{\partial N}{\partial P_i}D}{D^2 + N^2} \tag{112}$$

The partial derivates of the numerator and the denominator with respect to P_i can be obtained as follows:

$$\frac{\partial N}{\partial P_i} = \frac{\partial}{\partial P_i}\left\{\frac{2}{n}\sum\sum P_j P_k \cos Q_j \sin Q_k\right\} - \frac{\partial}{\partial P_i}\{\sum P_j^2 \sin 2Q_j\} \tag{113}$$

$$= -2P_i \sin 2Q_i + \frac{2}{n}\frac{\partial}{\partial P_i}\{P_1^2 c_1 s_1 + P_1 c_1 P_2 s_2 + P_1 c_1 P_3 s_3 + P_1 c_1 P_4 s_4 + \dots$$

$$+ P_2 c_2 P_1 s_1 + P_2^2 c_2 s_2 + P_2 c_2 P_3 s_3 + P_2 c_2 P_4 s_4 + \dots$$

$$+ P_3 c_3 P_1 s_1 + P_3 c_3 P_2 s_2 + P_3^2 c_3 s_3 + P_3 c_3 P_4 s_4 + \dots\} \tag{114}$$

[†] See [ARRAS97] for the results with weights. Performance comparison results of three different ways to determine α and r are also briefly given.

$$= \frac{2}{n} \left(\frac{\partial}{\partial P_i} \sum \{P_j \cos Q_j P_i \sin Q_i\} + \frac{\partial}{\partial P_i} \sum \{P_i \cos Q_i P_j \sin Q_j\} \right) - 2P_i \sin 2Q_i \tag{115}$$

$$= \frac{2}{n} \left(\sin Q_i \sum P_j \cos Q_j + \cos Q_i \sum P_j \sin Q_j \right) - 2P_i \sin 2Q_i \tag{116}$$

$$= \frac{2}{n} \left(\sin Q_i n\bar{x} + \cos Q_i n\bar{y} \right) - 2P_i \sin 2Q_i \tag{117}$$

$$= 2(\bar{x} \sin Q_i + \bar{y} \cos Q_i - P_i \sin 2Q_i) \tag{118}$$

The mean values $\bar{x}$ and $\bar{y}$ are those of equation (111).

$$\frac{\partial D}{\partial P_i} = \frac{\partial}{\partial P_i} \left\{ \frac{1}{n} \sum \sum P_j P_k \cos(Q_i + Q_j) \right\} - \frac{\partial}{\partial P_i} \{ \sum P_j^2 \cos 2Q_j \} \tag{119}$$

$$= -2P_i \cos 2Q_i + \frac{1}{n} \frac{\partial}{\partial P_i} \{ P_1^2 c_{1+1} + P_1 P_2 c_{1+2} + P_1 P_3 c_{1+3} + P_1 P_4 c_{1+4} + \dots$$
$$+ P_2 P_1 c_{2+1} + P_2^2 c_{2+2} + P_2 P_3 c_{2+3} + P_2 P_4 c_{2+4} + \dots$$
$$+ P_3 P_1 c_{3+1} + P_3 P_2 c_{3+2} + P_3^2 c_{3+3} + P_3 P_4 c_{3+4} + \dots \} \tag{120}$$

$$= \frac{1}{n} \left[\frac{\partial}{\partial P_i} \{ \sum P_j P_i \cos(Q_i + Q_j) \} + \frac{\partial}{\partial P_i} \{ \sum P_i P_j \cos(Q_j + Q_i) \} \right] - 2P_i \cos 2Q_i \tag{121}$$

$$= \frac{2}{n} \frac{\partial}{\partial P_i} \{ \sum P_j P_i \cos(Q_i + Q_j) \} - 2P_i \cos 2Q_i \tag{122}$$

$$= \frac{2}{n} \sum P_j \cos(Q_i + Q_j) - 2P_i \cos 2Q_i \tag{123}$$

$$= \frac{2}{n} \sum P_j \cos Q_i \cos Q_j - \frac{2}{n} \sum P_j \sin Q_i \sin Q_j - 2P_i \cos 2Q_i \tag{124}$$

$$= \frac{2}{n} \cos Q_i \sum P_j \cos Q_j - \frac{2}{n} \sin Q_i \sum P_j \sin Q_j - 2P_i \cos 2Q_i \tag{125}$$

$$= \frac{2}{n} \cos Q_i n\bar{x} - \frac{2}{n} \sin Q_i n\bar{y} - 2P_i \cos 2Q_i \tag{126}$$

$$= 2(\bar{x} \cos Q_i - \bar{y} \sin Q_i - P_i \cos 2Q_i) \tag{127}$$

Substituting into equation (112) gives

$$\frac{\partial \alpha}{\partial P_i} = \frac{1}{2} \cdot \frac{\frac{\partial D}{\partial P_i} N - \frac{\partial N}{\partial P_i} D}{D^2 + N^2} \tag{128}$$

$$= \frac{1}{2} \cdot \frac{2(\bar{x} \cos Q_i - \bar{y} \sin Q_i - P_i \cos 2Q_i)N - 2(\bar{x} \sin Q_i + \bar{y} \cos Q_i - P_i \sin 2Q_i)D}{D^2 + N^2} \tag{129}$$

$$= \frac{N(\bar{x} \cos Q_i - \bar{y} \sin Q_i - P_i \cos 2Q_i) - D(\bar{x} \sin Q_i + \bar{y} \cos Q_i - P_i \sin 2Q_i)}{D^2 + N^2} \tag{130}$$

It remains the derivate of r:

$$\frac{\partial r}{\partial P_i} = \frac{\partial}{\partial P_i} \left\{ \frac{1}{n} \sum P_j \cos(Q_j - \alpha) \right\} \tag{131}$$

$$= \frac{1}{n} \sum \frac{\partial}{\partial P_i} \{ P_j \cos(Q_j - \alpha) \} \tag{132}$$

$$= \frac{1}{n}\sum \cos(Q_j - \alpha)\frac{\partial}{\partial P_i}\{P_j\} + \frac{1}{n}\sum P_j\frac{\partial}{\partial P_i}\{\cos(Q_j - \alpha)\} \tag{133}$$

$$= \frac{1}{n}\cos(Q_i - \alpha) + \frac{1}{n}\sum P_j\sin(Q_j - \alpha)\frac{\partial\alpha}{\partial P_i} \tag{134}$$

$$= \frac{1}{n}\cos(Q_i - \alpha) + \frac{\partial\alpha}{\partial P_i}\frac{1}{n}\sum P_j\sin(Q_j - \alpha) \tag{135}$$

$$= \frac{1}{n}\cos(Q_i - \alpha) + \frac{\partial\alpha}{\partial P_i}\cdot\frac{\partial r}{\partial\alpha} \tag{136}$$

$$= \frac{1}{n}\cos(Q_i - \alpha) + \frac{\partial\alpha}{\partial P_i}(\bar{y}\cos\alpha - \bar{x}\sin\alpha) \tag{137}$$

Note that we made use of the already known expression $\partial\alpha/\partial P_i$.

Let us summarize the results. The sought covariance matrix is

$$C_{AR} = \begin{bmatrix} \sigma_A^2 & \sigma_{AR} \\ \sigma_{AR} & \sigma_R^2 \end{bmatrix} \tag{138}$$

with elements

$$\sigma_A^2 = \frac{1}{(D^2 + N^2)^2}\sum [N(\bar{x}\cos\theta_i - \bar{y}\sin\theta_i - \rho_i\cos 2\theta_i) - D(\bar{x}\sin\theta_i + \bar{y}\cos\theta_i - \rho_i\sin 2\theta_i)]^2\sigma_{\rho_i}^2 \tag{139}$$

$$\sigma_R^2 = \sum \left[\frac{1}{n}\cos(\theta_i - \alpha) + \frac{\partial\alpha}{\partial P_i}(\bar{y}\cos\alpha - \bar{x}\sin\alpha)\right]^2\sigma_{\rho_i}^2 \tag{140}$$

$$\sigma_{AR} = \sum \frac{\partial\alpha}{\partial P_i}\frac{\partial r}{\partial P_i}\sigma_{\rho_i}^2 \tag{141}$$

All elements of C_{AR} are of complexity $O(n)$ and allow extensive reuse of already computed expressions. This makes them suitable for fast calculation under real time conditions and limited computing power.