Australian Government
Department of Defence
Defence Science and
Technology Organisation

Proofs and Techniques Useful for Deriving the Kalman Filter

Don Koks

Electronic Warfare and Radar Division

Defence Science and Technology Organisation

DSTO–TN–0803

ABSTRACT

This note is a tutorial in matrix manipulation and the normal distribution of statistics, concepts that are important for deriving and analysing the Kalman Filter, a basic tool of signal processing. We focus on the proof of the well-known fact that the sum of two n-dimensional normal probability density functions is also normal. While this theorem is usually taken for granted in the signal processing field, proving it provides an insightful excursion into techniques such as Gaussian integrals and the Matrix Inversion Lemma.

Published by

Defence Science and Technology Organisation
PO Box 1500
Edinburgh, SA 5111, Australia

Telephone: (08) 8259 5555
Facsimile: (08) 8259 6567

AR No. AR–014-097
February, 2008

APPROVED FOR PUBLIC RELEASE

Proofs and Techniques Useful for Deriving the Kalman Filter

EXECUTIVE SUMMARY

Much analysis in the field of tracking and signal processing involves many parameters. One important example is the Kalman Filter, an algorithm that updates the estimated values of parameters based on their previous estimated values and a set of observations.

Parameters such as these are usually best arranged in a vector for economy of language. Any linearity inherent in the technique being described can then be expressed using matrix language. Deriving and analysing the Kalman Filter is one such example of this, so that a good command of matrix manipulation becomes useful to the field. For example, matrices and vectors are used to manipulate the normal probability density functions used in the Kalman Filter.

In this note, we have used some of these techniques to prove the well-known fact that the sum of two n-dimensional normal density functions is also normal. While this theorem is usually taken for granted in the signal processing field, proving it is an insightful exercise in applying some useful matrix techniques, such as Gaussian integrals and the Matrix Inversion Lemma.

Author

Don Koks
Electronic Warfare and Radar Division

Don Koks completed a doctorate in mathematical physics at Adelaide University in 1996, with a dissertation describing the use of quantum statistical methods to analyse decoherence, entropy and thermal radiance in both the early universe and black hole theory. He holds a Bachelor of Science from the University of Auckland in pure and applied mathematics, and a Master of Science in physics from the same university with a thesis in applied accelerator physics (proton-induced X ray and γ ray emission for trace element analysis). He has worked on the accelerator mass spectrometry programme at the Australian National University in Canberra, as well as in commercial internet development.

Currently he is a Research Scientist with the Maritime Systems group in the Electronic Warfare and Radar Division at DSTO, specialising in jamming, three-dimensional rotations, and geolocation. He has published a book on mathematical physics called *Explorations in Mathematical Physics: the Concepts Behind an Elegant Language* (Springer, 2006).

Contents

Appendices

1 Getting Started: the Proof for One-Dimensional Variables

Much analysis in the field of tracking and signal processing involves many parameters. One important example is the Kalman Filter, an algorithm that updates the estimated values of parameters based on their previous estimated values and a set of observations.

Parameters such as these are usually best arranged in a vector for economy of language. Any linearity inherent in the technique being described can then be expressed using matrix language. Deriving and analysing the Kalman Filter is one such example of this, so that a good command of matrix manipulation becomes useful to the field. For example, matrices and vectors are used to manipulate the normal probability density functions used in the Kalman Filter.

In this note, we have used some of these techniques to prove the well-known fact that the sum of two n-dimensional normal density functions is also normal. While this theorem is usually taken for granted in the signal processing field, proving it is an insightful exercise in applying some useful matrix techniques, such as Gaussian integrals and the Matrix Inversion Lemma.

We begin by stating the theorem to be proved: that the sum of two Gaussian density functions is another Gaussian function. Its mean is the sum of the individual means, and its variance (or covariance in the n-dimensional case) is the sum of the individual variances (or covariances). The proof of this fact uses some techniques and results that are useful knowledge for anyone undertaking analytical work in the field of tracking. These techniques and proofs are, in fact, not easy to locate in the literature, and so we present them here. We have not aimed for any extreme economy in how the process has been carried out. Rather, the calculation is done from a first-principles point of view, precisely because of its effectiveness as an exercise in matrix manipulation.

Two results that are needed are given in the appendices. The first is the result of an n-dimensional integration of a Gaussian function. The second appendix gives a conveniently short form of the very useful Matrix Inversion Lemma, from which all other forms of that lemma can be derived in a straightforward way (as demonstrated by an example in that appendix).

The sum-of-Gaussians result is first proved here in one dimension, to give a feel for the approach to be followed in the n-dimensional case. Consider two random variables

$$x \sim \mathcal{N}(\bar{x},\ \sigma_x^2) \quad \text{and } y \sim \mathcal{N}(\bar{y},\ \sigma_y^2)\,, \quad \text{with } z \equiv x + y\,, \tag{1.1}$$

by which we mean there are two Gaussian functions being considered:

$$p_x(x) \equiv \frac{1}{\sigma_x\sqrt{2\pi}} \exp\frac{-(x-\bar{x})^2}{2\sigma_x^2}\,, \quad p_y(y) \equiv \frac{1}{\sigma_y\sqrt{2\pi}} \exp\frac{-(y-\bar{y})^2}{2\sigma_y^2}\,. \tag{1.2}$$

The task is to compute the sum density, $p(z)$. If x, y are independent, then the probability $p(z)\,\mathrm{d}z$ that z is found in some interval $[z,\ z+\mathrm{d}z]$ equals the product of the probabilities that x is found in the interval $[x,\ x+\mathrm{d}x]$, and y is found in a corresponding interval constrained to ensure that $y = z - x$:

$$p(z)\,\mathrm{d}z = \int_x p_x(x)\mathrm{d}x\; p_y(y)\mathrm{d}y\big|_{y=z-x}\,. \tag{1.3}$$

Here are two different ways to analyse this integral.

First way: change of variables Consider new variables X, z, functions of x, y defined via

$$x = X\,, \quad y = z - X\,. \tag{1.4}$$

Changing variables in (1.3) gives [1]

$$\begin{aligned} p(z)\,\mathrm{d}z &= \int_X p_x(X)\,p_y(z-X)\left|\frac{\partial(x,y)}{\partial(X,z)}\right|\mathrm{d}X\,\mathrm{d}z \\ &= \int_X p_x(X)\,p_y(z-X)\,\mathrm{d}X\,\mathrm{d}z\,, \qquad \text{since } \frac{\partial(x,y)}{\partial(X,z)} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1\,. \end{aligned} \tag{1.5}$$

But X is now just a dummy variable of integration, so change it to x to give the required expression:

$$p(z) = \int p_x(x)\,p_y(z-x)\,\mathrm{d}x\,. \tag{1.6}$$

Second way: graphical viewpoint Alternatively, refer to Fig. 1, the blue region of which shows the points (x, y) such that y is constrained to an infinitesimal region around $y = z - x$, and z lies in $[z,\ z + \mathrm{d}z]$. The area of the shaded tile is $\mathrm{d}x\,\mathrm{d}y$. But this area is also $\mathrm{d}x\,\mathrm{d}z$. Thus (1.3) becomes

$$p(z)\,\mathrm{d}z = \int_x p_x(x)\,p_y(z-x)\,\mathrm{d}x\,\mathrm{d}z\,, \tag{1.7}$$

in which case

$$p(z) = \int p_x(x)\,p_y(z-x)\,\mathrm{d}x\,, \tag{1.8}$$

agreeing with (1.6).

Equation (1.6) is a convolution integral, and relates the technique of convolution to a summing of random variables. Using it, we are able to construct $p(z)$ given the two functions in (1.2):

$$p(z) = \frac{1}{\sigma_x\sigma_y 2\pi}\int \exp\left[\frac{-(x-\bar{x})^2}{2\sigma_x^2} - \frac{(z-x-\bar{y})^2}{2\sigma_y^2}\right]\mathrm{d}x\,. \tag{1.9}$$

The brackets of (1.9) expand to

$$-x^2\left(\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_y^2}\right) + x\left(\frac{\bar{x}}{\sigma_x^2} + \frac{z-\bar{y}}{\sigma_y^2}\right) - \frac{\bar{x}^2}{2\sigma_x^2} - \frac{(z-\bar{y})^2}{2\sigma_y^2}\,, \tag{1.10}$$

which, being a quadratic in x, allows the integral (1.9) to be done:

$$p(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}}\exp\left[\frac{\left(\frac{\bar{x}}{\sigma_x^2} + \frac{z-\bar{y}}{\sigma_y^2}\right)^2}{2\left(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2}\right)} - \frac{\bar{x}^2}{2\sigma_x^2} - \frac{(z-\bar{y})^2}{2\sigma_y^2}\right]. \tag{1.11}$$

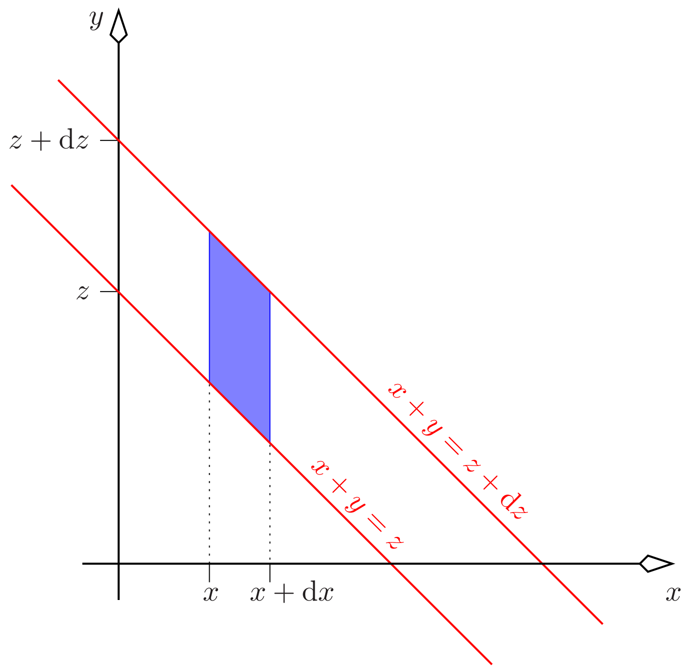

Figure 1: *A graphical depiction of the change of variables in* (1.7)

The brackets of (1.11) simplify considerably. Equation (1.11) can be written more succinctly by defining $\sigma \equiv \sqrt{\sigma_x^2 + \sigma_y^2}$, producing

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-[z - (\bar{x} + \bar{y})]^2}{2\sigma^2} . \tag{1.12}$$

But this is a Gaussian function with mean $\bar{x} + \bar{y}$ and variance $\sigma_x^2 + \sigma_y^2$. That is,

$$x + y \sim \mathcal{N}(\bar{x} + \bar{y},\ \sigma_x^2 + \sigma_y^2) , \tag{1.13}$$

as was required to be proved.

2 The Proof for n-Dimensional Variables

The proof that the sum of two n-dimensional Gaussians gives another Gaussian follows the same line of reasoning as in the 1-dimensional case, but is more involved owing to the many matrix manipulations required.

Begin with two n-dimensional Gaussian variables (all vectors are columns in what follows):

$$\boldsymbol{x} = [x_1 \ldots x_n]^t , \quad \boldsymbol{y} = [y_1 \ldots y_n]^t , \tag{2.1}$$

with

$$\boldsymbol{x} \sim \mathcal{N}(\bar{\boldsymbol{x}},\ P_x) \quad \text{and} \quad \boldsymbol{y} \sim \mathcal{N}(\bar{\boldsymbol{y}},\ P_y) . \tag{2.2}$$

Their density functions are extensions of (1.2):

$$p_x(\boldsymbol{x}) \equiv \frac{1}{(2\pi)^{n/2}\ |P_x|^{1/2}}\ \exp\frac{-1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^t P_x^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}})\,,$$
$$p_y(\boldsymbol{y}) \equiv \frac{1}{(2\pi)^{n/2}\ |P_y|^{1/2}}\ \exp\frac{-1}{2}(\boldsymbol{y}-\bar{\boldsymbol{y}})^t P_y^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}})\,. \tag{2.3}$$

Define $\boldsymbol{z} \equiv \boldsymbol{x}+\boldsymbol{y}$. We wish to show that $\boldsymbol{z}$ is normally distributed with mean $\bar{\boldsymbol{x}}+\bar{\boldsymbol{y}}$ and covariance $P_x + P_y$.

The proof begins by creating a convolution integral, just as in the 1-dimensional case. To see how it comes about, consider (for brevity) the case of $n = 2$ dimensions. Just as $\mathrm{d}y = \mathrm{d}z$ in the 1-dimensional case of Fig. 1, so also in the 2-dimensional case we have

$$\begin{aligned} p(\boldsymbol{z})\,\mathrm{d}z_1\,\mathrm{d}z_2 &= \text{probability that } x_1 \in [x_1,\ x_1+\mathrm{d}x_1] \text{ and } x_2 \in [x_2,\ x_2+\mathrm{d}x_2] \\ &\qquad \text{and } y_1 \in [y_1,\ y_1+\mathrm{d}z_1] \text{ and } y_2 \in [y_2,\ y_2+\mathrm{d}z_2] \\ &= \int_{x_1}\int_{x_2} p_x(x_1,x_2)\ p_y(z_1-x_1,\ z_2-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}z_1\,\mathrm{d}z_2\,, \end{aligned} \tag{2.4}$$

so that

$$p(\boldsymbol{z}) = \iint p_x(x_1,x_2)\ p_y(z_1-x_1,\ z_2-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\,. \tag{2.5}$$

This is seen to extend to n dimensions, in which case the required integral is

$$p(\boldsymbol{z}) = \frac{1}{(2\pi)^n\ |P_xP_y|^{1/2}}\int \exp\frac{-1}{2}\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})^t P_x^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{z}-\boldsymbol{x}-\bar{\boldsymbol{y}})^t P_y^{-1}(\boldsymbol{z}-\boldsymbol{x}-\bar{\boldsymbol{y}})\right]\mathrm{d}x_1\ldots\mathrm{d}x_n. \tag{2.6}$$

The integration is over $\boldsymbol{x}$, so collecting terms in $\boldsymbol{x}$ within the brackets in (2.6) gives

$$\begin{aligned} p(\boldsymbol{z}) = \frac{1}{(2\pi)^n\ |P_xP_y|^{1/2}}\int \exp\Bigg[&-\boldsymbol{x}^t\frac{P_x^{-1}+P_y^{-1}}{2}\boldsymbol{x} + [\bar{\boldsymbol{x}}^t P_x^{-1} + (\boldsymbol{z}-\bar{\boldsymbol{y}})^t P_y^{-1}]\boldsymbol{x} \\ &-\frac{1}{2}\bar{\boldsymbol{x}}^t P_x^{-1}\bar{\boldsymbol{x}} - \frac{1}{2}(\boldsymbol{z}-\bar{\boldsymbol{y}})^t P_y^{-1}(\boldsymbol{z}-\bar{\boldsymbol{y}})\Bigg]\,\mathrm{d}x_1\ldots\mathrm{d}x_n\,, \end{aligned} \tag{2.7}$$

which integrates via (A2) to give

$$\begin{aligned} p(\boldsymbol{z}) = &\frac{\pi^{n/2}}{(2\pi)^n\ |P_xP_y|^{1/2}\ \left|\frac{P_x^{-1}+P_y^{-1}}{2}\right|^{1/2}} \times \\ &\exp\Bigg[\frac{1}{4}\left(\bar{\boldsymbol{x}}^t P_x^{-1} + (\boldsymbol{z}-\bar{\boldsymbol{y}})^t P_y^{-1}\right)\left(\frac{P_x^{-1}+P_y^{-1}}{2}\right)^{-1}\left(P_x^{-1}\bar{\boldsymbol{x}} + P_y^{-1}(\boldsymbol{z}-\bar{\boldsymbol{y}})\right) \\ &\qquad -\frac{1}{2}\bar{\boldsymbol{x}}^t P_x^{-1}\bar{\boldsymbol{x}} - \frac{1}{2}(\boldsymbol{z}-\bar{\boldsymbol{y}})^t P_y^{-1}(\boldsymbol{z}-\bar{\boldsymbol{y}})\Bigg]\,. \end{aligned} \tag{2.8}$$

Define a matrix P such that $P^{-1} \equiv P_x^{-1} + P_y^{-1}$. In that case

$$\begin{aligned} \left|\frac{P_x^{-1}+P_y^{-1}}{2}\right|^{1/2} &= \left|\frac{P^{-1}}{2}\right|^{1/2} = \frac{|P|^{-1/2}}{2^{n/2}}\,, \quad \text{and} \\ \left(\frac{P_x^{-1}+P_y^{-1}}{2}\right)^{-1} &= \left(\frac{P^{-1}}{2}\right)^{-1} = 2P\,. \end{aligned} \tag{2.9}$$

Thus

$$p(\boldsymbol{z}) = \frac{|P|^{1/2}}{(2\pi)^{n/2}\,|P_x P_y|^{1/2}}\ \exp B\,, \tag{2.10}$$

where (setting $\boldsymbol{\alpha} \equiv \boldsymbol{z} - \bar{\boldsymbol{y}}$)

$$\begin{aligned} 2B &= \left(\bar{\boldsymbol{x}}^t P_x^{-1} + \boldsymbol{\alpha}^t P_y^{-1}\right) P \left(P_x^{-1}\bar{\boldsymbol{x}} + P_y^{-1}\boldsymbol{\alpha}\right) - \bar{\boldsymbol{x}}^t P_x^{-1}\bar{\boldsymbol{x}} - \boldsymbol{\alpha}^t P_y^{-1}\boldsymbol{\alpha} \\ &= \bar{\boldsymbol{x}}^t \left(P_x^{-1} P P_x^{-1} - P_x^{-1}\right)\bar{\boldsymbol{x}} + 2\bar{\boldsymbol{x}}^t P_x^{-1} P P_y^{-1}\boldsymbol{\alpha} + \boldsymbol{\alpha}^t \left(P_y^{-1} P P_y^{-1} - P_y^{-1}\right)\boldsymbol{\alpha}\,. \end{aligned} \tag{2.11}$$

There are three expressions involving P, P_x, P_y in the last line of (2.11) that need simplifying. This can be done using the Matrix Inversion Lemma, explained in Appendix B. Hoping to prove that the covariance of $\boldsymbol{z}$ is $P_x + P_y$, we will aim to have $P_x + P_y$ appear wherever possible.

$$\begin{aligned} P_x^{-1} P P_x^{-1} - P_x^{-1} &\overset{\text{(B4)}}{=\!=} -(P_x + P_y)^{-1}\,, \\ P_x^{-1} P P_y^{-1} &= [P_y(P_x^{-1} + P_y^{-1})P_x]^{-1} = (P_x + P_y)^{-1}\,, \\ P_y^{-1} P P_y^{-1} - P_y^{-1} &= -(P_x + P_y)^{-1} \quad \text{(from two lines up with } x \leftrightarrow y\text{).} \end{aligned} \tag{2.12}$$

Thus

$$\begin{aligned} 2B &= -\bar{\boldsymbol{x}}^t (P_x + P_y)^{-1}\bar{\boldsymbol{x}} + 2\bar{\boldsymbol{x}}^t (P_x + P_y)^{-1}\boldsymbol{\alpha} - \boldsymbol{\alpha}^t (P_x + P_y)^{-1}\boldsymbol{\alpha} \\ &= -(\bar{\boldsymbol{x}} - \boldsymbol{\alpha})^t (P_x + P_y)^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\alpha}) \\ &= -[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})]^t (P_x + P_y)^{-1}[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})]\,. \end{aligned} \tag{2.13}$$

Finally, (2.10) can be written as

$$\begin{aligned} p(\boldsymbol{z}) &= \frac{1}{(2\pi)^{n/2}\,|P_x^{-1} + P_y^{-1}|^{1/2}\,|P_x P_y|^{1/2}}\ \exp \frac{-1}{2}[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})]^t (P_x + P_y)^{-1}[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})] \\ &= \frac{1}{(2\pi)^{n/2}\,|P_x + P_y|^{1/2}}\ \exp \frac{-1}{2}[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})]^t (P_x + P_y)^{-1}[\boldsymbol{z} - (\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}})]\,. \end{aligned} \tag{2.14}$$

That is, $\boldsymbol{z} \sim \mathcal{N}(\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}},\ P_x + P_y)$, as was required to be proved.

Acknowledgements

The author thanks Sanjeev Arulampalam for discussions while writing this paper.

References

1. D. Koks (2006), *Explorations in Mathematical Physics*, Springer New York. This covers multidimensional integration in some detail.

Appendix A Calculating an n-Dimensional Gaussian Integral

This appendix takes the well-known 1-dimensional result

$$\int_{-\infty}^{\infty} \mathrm{e}^{-ax^2+bx}\mathrm{d}x = \sqrt{\frac{\pi}{a}}\,\exp\frac{b^2}{4a} \tag{A1}$$

and generalises it to the less well-known but very useful n-dimensional result:

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{x}^t A\boldsymbol{x}+\boldsymbol{b}^t\boldsymbol{x}}\,\mathrm{d}x_1\ldots\mathrm{d}x_n = \frac{\pi^{n/2}}{|A|^{1/2}}\,\exp\frac{1}{4}\boldsymbol{b}^t A^{-1}\boldsymbol{b}\,, \tag{A2}$$

where A is a real symmetric $n\times n$ matrix, and $\boldsymbol{b}, \boldsymbol{x}$ (and $\boldsymbol{u}$ in what follows) are n-dimensional column vectors. The "t" superscript denotes the matrix transpose.

First, note that because A is real and symmetric, it can be orthogonally diagonalised to give $A = PA'P^t$ with $P^{-1} = P^t$ and A' diagonal. Use this P to create a change of variables to $\boldsymbol{u}$, via $\boldsymbol{x} = P\boldsymbol{u}$. Denote the left-hand side of (A2) by I, which we must show equals the right-hand side of (A2). The change of variables converts the left-hand side of (A2) to [1]

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{u}^t A'\boldsymbol{u}+\boldsymbol{b}^t P\boldsymbol{u}} \left|\frac{\partial(x_1,\ldots,x_n)}{\partial(u_1,\ldots,u_n)}\right| \mathrm{d}u_1\ldots\mathrm{d}u_n\,. \tag{A3}$$

Since the elements of P are constants, the ij^{th} element of the Jacobian matrix is P_{ij}. Thus the Jacobian matrix is just P, so that the Jacobian determinant is

$$\frac{\partial(x_1,\ldots,x_n)}{\partial(u_1,\ldots,u_n)} = |P| \in \{\pm 1\}\,, \quad \text{so that } \left|\frac{\partial(x_1,\ldots,x_n)}{\partial(u_1,\ldots,u_n)}\right| = 1\,. \tag{A4}$$

Now set $\boldsymbol{b}'^t \equiv \boldsymbol{b}^t P$, and write I as

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{u}^t A'\boldsymbol{u}+\boldsymbol{b}'^t\boldsymbol{u}}\,\mathrm{d}u_1\ldots\mathrm{d}u_n = \int_{-\infty}^{\infty} \exp\Big[\sum_i -A'_{ii}u_i^2 + b'_i u_i\Big]\mathrm{d}u_1\ldots\mathrm{d}u_n \\ &= \prod_i \int_{-\infty}^{\infty} \exp\Big[-A'_{ii}u_i^2 + b'_i u_i\Big]\mathrm{d}u_1\ldots\mathrm{d}u_n \overset{\text{(A1)}}{=\!=} \prod_i \sqrt{\frac{\pi}{A'_{ii}}}\,\exp\frac{{b'_i}^2}{4A'_{ii}}\,. \end{aligned} \tag{A5}$$

But $\prod_i A'_{ii} = |A'| = |P^t AP| = |A|$, so

$$I = \frac{\pi^{n/2}}{|A|^{1/2}}\,\exp\frac{1}{4}\sum_i \frac{{b'_i}^2}{A'_{ii}}\,. \tag{A6}$$

Also

$$\sum_i \frac{{b'_i}^2}{A'_{ii}} = \boldsymbol{b}'^t \begin{bmatrix} 1/A'_{11} & & 0 \\ & \ddots & \\ 0 & & 1/A'_{nn} \end{bmatrix} \boldsymbol{b}' = \boldsymbol{b}'^t A'^{-1}\boldsymbol{b}' = \boldsymbol{b}^t PP^t A^{-1} PP^t \boldsymbol{b} = \boldsymbol{b}^t A^{-1}\boldsymbol{b}\,. \tag{A7}$$

Thus (A6) becomes

$$I = \frac{\pi^{n/2}}{|A|^{1/2}}\,\exp\frac{1}{4}\boldsymbol{b}^t A^{-1}\boldsymbol{b}\,, \tag{A8}$$

which is the right-hand side of (A2). QED.

Appendix B Matrix Inversion Lemma

The Matrix Inversion Lemma is often used and very powerful in the matrix analysis of signal processing. It comes in various forms, but they are all easily derived from the following basic form of the lemma. For any matrices A and B not necessarily square, as long as the products AB and BA exist and the relevant matrices are invertible,

$$(I + AB)^{-1} = I - A(I + BA)^{-1}B\,, \tag{B1}$$

where by "I" throughout this appendix we mean the identity matrix of size appropriate to its use.

The lemma can be proved by multiplying the inverse of the left-hand side of (B1) by its right-hand side and inspecting the result:

$$\begin{aligned}\mathrm{LHS}^{-1}\,.\,\mathrm{RHS} &= (I + AB)\big[I - A(I + BA)^{-1}B\big]\\ &= I + AB - A(I + BA)^{-1}B - ABA(I + BA)^{-1}B\\ &= I + A\big[I - (I + BA)^{-1} - BA(I + BA)^{-1}\big]B\\ &= I + A\big[I - (I + BA)(I + BA)^{-1}\big]B\\ &= I\,. \end{aligned} \tag{B2}$$

In that case, LHS = RHS, and the lemma is proved.

More complicated versions of the lemma make good use of the fact that $(PQ)^{-1} = Q^{-1}P^{-1}$ for any invertible matrices P, Q. For example, apply the lemma to $(A + BCD)^{-1}$:

$$\begin{aligned}(A + BCD)^{-1} &= \big[(I + BCDA^{-1})A\big]^{-1}\\ &= A^{-1}(I + BCDA^{-1})^{-1}\\ &\overset{\text{(B1)}}{=\!=} A^{-1}\big(I - BC[I + DA^{-1}BC]^{-1}DA^{-1}\big)\\ &= A^{-1}\big(I - B\big[(I + DA^{-1}BC)C^{-1}\big]^{-1}DA^{-1}\big)\\ &= A^{-1}\big(I - B\big[C^{-1} + DA^{-1}B\big]^{-1}DA^{-1}\big)\,. \end{aligned} \tag{B3}$$

Finally,

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}\,, \tag{B4}$$

which is a common form of the Matrix Inversion Lemma.

Page classification:UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA				1. CAVEAT/PRIVACY MARKING

2. TITLE	3. SECURITY CLASSIFICATION
Proofs and Techniques Useful for Deriving the Kalman Filter	Document (U) Title (U) Abstract (U)
4. AUTHOR	5. CORPORATE AUTHOR
Don Koks	Defence Science and Technology Organisation PO Box 1500 Edinburgh, SA 5111, Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–TN–0803	AR–014-097	Technical Note	February, 2008

8. FILE NUMBER	9. TASK NUMBER	10. SPONSOR	11. No. OF PAGES	12. No OF REFS
2007/1027487	NAV 05/222	DGMD	9	1

13. URL OF ELECTRONIC VERSION	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO–TN–0803.pdf	Chief, Electronic Warfare and Radar Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DSTO RESEARCH LIBRARY THESAURUS

Kalman filters	Tracking
Methodology	Proving
Analysis	Matrices
Signal processing	

19. ABSTRACT

This note is a tutorial in matrix manipulation and the normal distribution of statistics, concepts that are important for deriving and analysing the Kalman Filter, a basic tool of signal processing. We focus on the proof of the well-known fact that the sum of two n-dimensional normal probability density functions is also normal. While this theorem is usually taken for granted in the signal processing field, proving it provides an insightful excursion into techniques such as Gaussian integrals and the Matrix Inversion Lemma.

Page classification:UNCLASSIFIED