PROBABILISTIC ROBOTICS:
GAUSSIAN FILTERS

Pierre-Paul TACHER

1

1.1. We can define the state vector to include position and velocity:

$$Y_t = \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix}$$

1.2. Evolution of the state will be approximated to the second and first order, for x and $\dot{x}$ respectively :

$$\begin{aligned} x_{t+1} &= x_t + \Delta t \times \dot{x}_t + \frac{1}{2}(\Delta t)^2 \times \ddot{x}_t + o(\Delta t^2) \\ \Rightarrow x_{t+1} &\approx x_t + 1 \times \dot{x}_t + \frac{1}{2} \times \ddot{x}_t \end{aligned}$$

$$\begin{aligned} \dot{x}_{t+1} &= \dot{x}_t + \Delta t \times \ddot{x}_t + o(\Delta t) \\ \Rightarrow \dot{x}_{t+1} &\approx \dot{x}_t + \ddot{x}_t \end{aligned}$$

We can modelize evolution of the state with matrix equality:

$$\begin{bmatrix} x_{t+1} \\ \dot{x}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{1}{2}\ddot{x}_t \\ \ddot{x}_t \end{bmatrix}}_{\epsilon_t}$$

$$\begin{bmatrix} x_{t+1} \\ \dot{x}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}}_{G} \ddot{x}_t$$

We have

$$G \times G^T = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$

As a linear function of gaussian $\ddot{x}_t$, the random variable ϵ_t is known to be multivariate Gaussian $\mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \underbrace{\sigma^2 GG^T}_{R})$. The conditional law of random variable $\begin{bmatrix} x_{t+1} \\ \dot{x}_{t+1} \end{bmatrix}$ given $(\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix})$ is then known to be also Gaussian $\mathcal{N}(A \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix}, R)$.

Moreover, if we assume the random variable $\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix}$ to be Gaussian $\mathcal{N}(\mu_t, \Sigma_t)$, then $\begin{bmatrix} x_{t+1} \\ \dot{x}_{t+1} \end{bmatrix}$ is Gaussian $\mathcal{N}(\underbrace{A\mu_t}_{\overline{\mu_{t+1}}}, \underbrace{A\Sigma_t A^T + R}_{\overline{\Sigma_{t+1}}})$.

Note that since we do not incorporate measurements for now, we have using the notation of the book

$$\overline{\mu_{t+1}} = \mu_{t+1}$$
$$\overline{\Sigma_{t+1}} = \Sigma_{t+1}$$

1.3. From the previous relations, it is clear that

$$\forall t \in \mathbb{N}, \mu_t = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let's t compute the first covariance matrices, using $\sigma^2 = 1$:

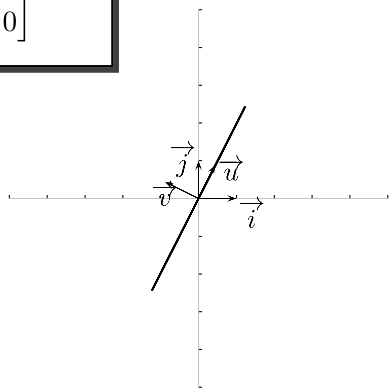
t	Σ_t
0	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
1	$\begin{bmatrix} 0.25 & 0.5 \\ 0.5 & 1 \end{bmatrix}$
2	$\begin{bmatrix} 2.5 & 2 \\ 2 & 2 \end{bmatrix}$
3	$\begin{bmatrix} 8.75 & 4.5 \\ 4.5 & 3 \end{bmatrix}$
4	$\begin{bmatrix} 21 & 8 \\ 8 & 4 \end{bmatrix}$
5	$\begin{bmatrix} 41.25 & 12.5 \\ 12.5 & 5 \end{bmatrix}$
6	$\begin{bmatrix} 71.5 & 18 \\ 18 & 6 \end{bmatrix}$
7	$\begin{bmatrix} 113.75 & 24.5 \\ 24.5 & 7 \end{bmatrix}$
8	$\begin{bmatrix} 170 & 32 \\ 32 & 8 \end{bmatrix}$
9	$\begin{bmatrix} 242.25 & 40.5 \\ 40.5 & 9 \end{bmatrix}$
10	$\begin{bmatrix} 332.5 & 50 \\ 50 & 10 \end{bmatrix}$

The uncertainty ellipses are centered at $\mu_t = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, have axes whose directions are given by the eigenvectors of Σ_t, and the semi-minor and semi-major axis are scaled by eigenvalues; for detailed derivation, cf. Σ_t is a symetric real matrix, it can be diagonalized in an orthonormal basis, with real eigenvalues

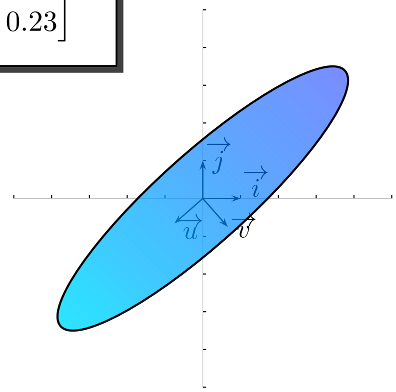
$$\Sigma_t = P_t D_t P_t^T$$

where P_t is an orthogonal matrix and D_t is diagonal. Note the first ellipse is degenerated since Σ_1 has a null eigenvalue.

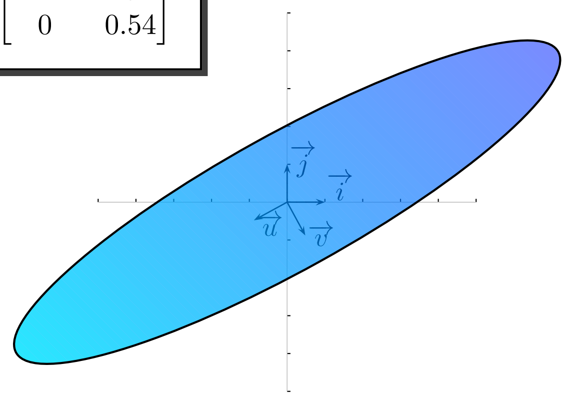
$$P = \begin{bmatrix} 0.45 & .89 \\ -0.89 & 0.45 \end{bmatrix}$$
$$D = \begin{bmatrix} 1.25 & 0 \\ 0 & 0 \end{bmatrix}$$



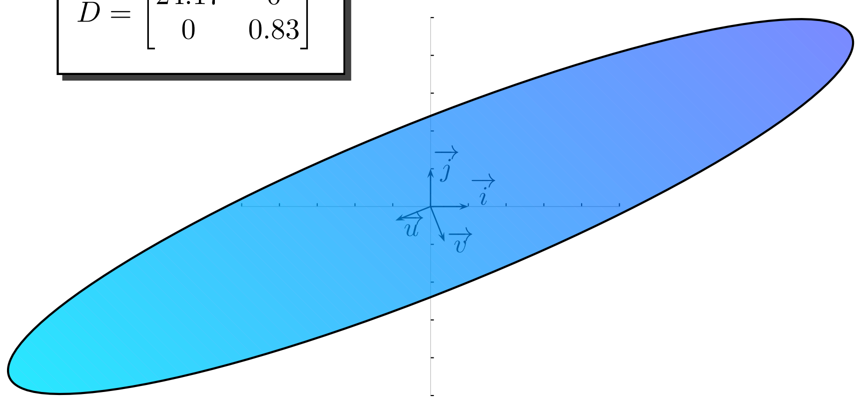
$$P = \begin{bmatrix} -0.75 & 0.66 \\ -0.66 & -0.75 \end{bmatrix}$$
$$D = \begin{bmatrix} 4.27 & 0 \\ 0 & 0.23 \end{bmatrix}$$



$$P = \begin{bmatrix} -0.88 & 0.48 \\ -0.48 & -0.88 \end{bmatrix}$$
$$D = \begin{bmatrix} 11.22 & 0 \\ 0 & 0.54 \end{bmatrix}$$



$$P = \begin{bmatrix} -0.93 & -0.37 \\ -0.37 & -0.93 \end{bmatrix}$$
$$D = \begin{bmatrix} 24.17 & 0 \\ 0 & 0.83 \end{bmatrix}$$



2

2.1. We can modelize the noisy observation of the position by:

$$\begin{aligned} z_t &= x_t + \xi_t \\ &= \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{C} \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + \xi_t \end{aligned}$$

$$P = \begin{bmatrix} -0.95 & -0.30 \\ -0.30 & -0.95 \end{bmatrix}$$
$$D = \begin{bmatrix} 45.14 & 0 \\ 0 & 1.11 \end{bmatrix}$$

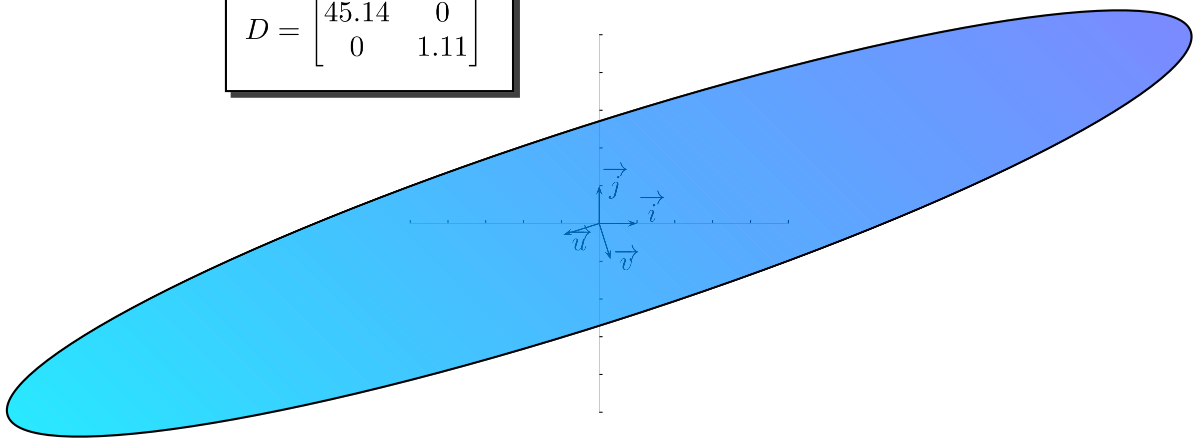


FIGURE 0. 0.95 uncertainty ellipses for the gaussian posteriors at dates $t \in [\![1, 5]\!]$

where the random variable ξ_t is gaussian $\xi_t \hookrightarrow \mathcal{N}(0, \underbrace{\sigma^2}_{Q})$.

2.2. Before the observation the moments are:

$$\bar{\mu}_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\bar{\Sigma}_5 = \begin{bmatrix} 41.25 & 12.5 \\ 12.5 & 5 \end{bmatrix}$$

The measurement update consists of

$$\begin{aligned} K_5 &= \bar{\Sigma}_5 C^T (C\bar{\Sigma}_5 C^T + Q)^{-1} \\ &= \begin{bmatrix} 0.80 \\ 0.24 \end{bmatrix} \\ \mu_5 &= \bar{\mu}_5 + K_5(z_5 - C\bar{\mu}_5) \\ &= \begin{bmatrix} 4.02 \\ 1.22 \end{bmatrix} \\ \Sigma_5 &= (I - K_5 C)\bar{\Sigma}_5 \\ &= \begin{bmatrix} 8.05 & 2.44 \\ 2.44 & 1.95 \end{bmatrix} \end{aligned}$$

We draw the corresponding ellipse:

3

Let's recall the definition of the characteristic function of a random variable: (a tad of theory of measure is needed)($\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^d$)

Let X be an random variable which takes it values in $\mathbb{R}^d$. The characteristic function of X is the function which goes from $\mathbb{R}^d$ to $\mathbb{C}$:

$$\begin{aligned} \forall t \in \mathbb{R}^d, \quad \varphi_X(t) &= E \exp i\langle X, t\rangle \\ &= \int_{\mathbb{R}^d} \exp i\langle x, t\rangle \, \mathrm{d}P_X(x) \end{aligned}$$

Note the characteristic function of X is the Fourier transform of its probability distribution. We

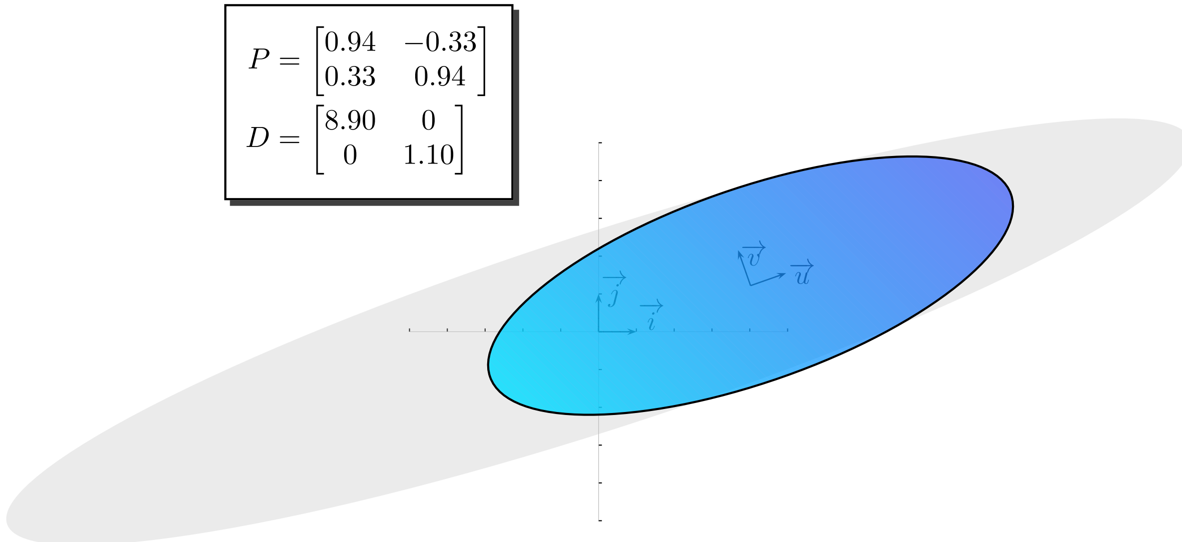


FIGURE 1. 0.95 uncertainty ellipse for the gaussian posterior after measurement of position at dates $t = 5$

will also use the following properties to derive the prediction update of Kalman filter:

Let two r.v. X_1 and X_2 be independant and take their values in $\mathbb{R}^d$. The characteristic function of their sum is given by

$$\forall t \in \mathbb{R}^d, \quad \varphi_{X_1+X_2}(t) = \varphi_{X_1}(t) \times \varphi_{X_2}(t)$$

Random variable X is Gaussian $X \hookrightarrow \mathcal{N}(\mu, \Lambda)$ if and only if its characteristic function φ_X is defined by

$$\forall t \in \mathbb{R}^d, \quad \varphi_X(t) = \exp i\langle \mu, t\rangle \exp[-\frac{1}{2}\langle \Lambda t, t\rangle]$$

We now restate the equation of prediction step for Kalman filter:

$$x_t = Ax_{t-1} + Bu_t + \epsilon_t$$

We assume x_{t-1} is Gaussian $x_{t-1} \hookrightarrow \mathcal{N}(\mu_{t-1}, \Sigma_{t-1})$, and ϵ_t is Gaussian $\epsilon_t \hookrightarrow \mathcal{N}(0, R_t)$. x_{t-1} and ϵ_t are assumed to be independant. u_t is deterministic. By linearity we know $y_t = Ax_{t-1} + Bu_t$ is Gaussian $y_t \hookrightarrow \mathcal{N}(A\mu_{t-1} + Bu_t, A\Sigma_{t-1}A^T)$. By the above property, we can compute the characteristic function of r.v. x_t

$$\begin{aligned} \forall u \in \mathbb{R}^d, \quad \varphi_{x_t}(u) &= \varphi_{y_t}(u) \times \varphi_{\epsilon_t}(u) \\ &= \exp i\langle A\mu_{t-1} + Bu_t, u\rangle \exp[-\frac{1}{2}\langle A\Sigma_{t-1}A^T u, u\rangle] \times \exp i\langle 0, u\rangle \exp[-\frac{1}{2}\langle R_t t, t\rangle] \\ &= \exp i\langle A\mu_{t-1} + Bu_t, u\rangle \exp[-\frac{1}{2}\langle (A\Sigma_{t-1}A^T + R_t)u, u\rangle] \end{aligned}$$

using again the property of gaussian in the other direction, this shows r.v. x_t is Gaussian $x_t \hookrightarrow \mathcal{N}(\underbrace{A\mu_{t-1} + Bu_t}_{\bar{\mu}_t}, \underbrace{A\Sigma_{t-1}A^T + R_t}_{\bar{\Sigma}_t})$.

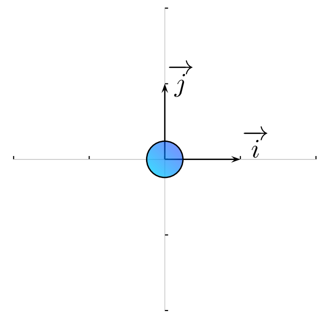


FIGURE 2. 0.95 uncertainty ellipses for the gaussian $\mathcal{N}(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0.01 & 0\\0 & 0.01\end{bmatrix})$

4

4.1. The system evolves according to

$$\begin{bmatrix}x'\\y'\\\theta\end{bmatrix} = \begin{bmatrix}x+\cos\theta\\y+\sin\theta\\\theta\end{bmatrix}$$

We will assume x, y and θ are gaussian r.v.:$x \hookrightarrow \mathcal{N}(0,\sigma_1^2)$ $y \hookrightarrow \mathcal{N}(0,\sigma_1^2)$ $\theta \hookrightarrow \mathcal{N}(0,\sigma_2^2)$, where $\sigma_1^2 = 0.01$ and $\sigma_2^2 = 10000$. We could compute the expectancy and variance of $\cos\theta$ and $\sin\theta$,but since the variability of θ is high with regards to 2π, and

$$\begin{aligned}\forall\theta\in\mathbb{R},\quad \cos\theta &= \cos(\theta - \left\lfloor\frac{\theta}{2\pi}\right\rfloor\times 2\pi)\\ \sin\theta &= \sin(\theta - \left\lfloor\frac{\theta}{2\pi}\right\rfloor\times 2\pi)\end{aligned}$$

it seems reasonable to assume $\alpha = \theta - \lfloor\frac{\theta}{2\pi}\rfloor \times 2\pi$ is an uniform r.v. $\alpha \hookrightarrow \mathcal{U}([0,2\pi])$. $\begin{bmatrix}x\\y\end{bmatrix}$ is gaussian $\begin{bmatrix}x\\y\end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & 0\\0 & \sigma_1^2\end{bmatrix})$. To draw the posterior, we will first draw the corresponding 0.95 uncertainty ellipse. We will use the following notation for this disk :

$$\mathcal{D} = \{z\in\mathbb{C}, |z| \leqslant r\}$$

Next we consider the function:

$$\begin{aligned} f : \mathcal{D}\times[0,2\pi] &\to \mathbb{C}\\ (z,\theta) &\mapsto z + e^{i\theta}\end{aligned}$$

We can show (cf.) that $f(\mathcal{D}\times[0,2\pi]) = \underbrace{\{z\in\mathbb{C}, 1-r \leqslant |z| \leqslant 1+r\}}_{\mathcal{E}}$. It can then be asserted r.v. $\begin{bmatrix}x'\\y'\end{bmatrix}$ belongs to $\mathcal{E}$ with a probability not lesser than 0.95.We represent $\mathcal{E}$ in.

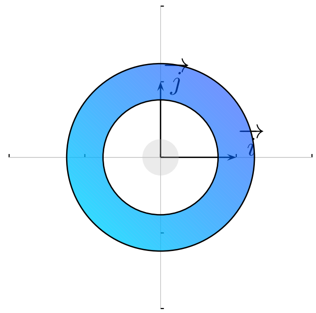


FIGURE 3. 0.95 uncertainty area for the random variable $\begin{bmatrix} x' \\ y' \end{bmatrix}$

4.2. In order to linearize the state evolution equation, we have to linearize $\cos\theta$ and $\sin\theta$ about $E\theta = 0$, which we know already makes no sense because of the high variability of θ.

$$\begin{aligned} \cos\theta &\underset{x\to 0}{=} 1 + o(\theta) \\ \Rightarrow \cos\theta &\approx 1 \\ \sin\theta &\underset{x\to 0}{=} \theta + o(\theta) \\ \Rightarrow \sin\theta &\approx \theta \end{aligned}$$

The motion model becomes

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{B}$$

If we suppose $\begin{bmatrix} x \\ y \\ \theta \end{bmatrix}$ is Gaussian $\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 10000 \end{bmatrix}}_{\Sigma})$, by linearity $\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix}$ is Gaussian $\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} \hookrightarrow \mathcal{N}(B, A\Sigma A^T)$. The marginal $\begin{bmatrix} x' \\ y' \end{bmatrix}$ is Gaussian (it results from the application of linear transformation, namely projection of the previous r.v.) $\begin{bmatrix} x \\ y \end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix} 01 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.01 & 0 \\ 0 & 10000 \end{bmatrix})$. From this follows the uncertainty ellipse in, which has been cropped for obvious reasons and unsurprisingly cannot capture the posterior correctly. Linearization makes sense only when we restrain the uncertainty on θ. Let's take $\sigma_2^2 = 1$, we have

$$\begin{bmatrix} x \\ y \end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.01 & 0 \\ 0 & 1.01 \end{bmatrix})$$

If $\sigma_2^2 = 0.5$,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.01 & 0 \\ 0 & 0.51 \end{bmatrix})$$

If $\sigma_2^2 = 0.25$,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} \hookrightarrow \mathcal{N}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.01 & 0 \\ 0 & 0.26 \end{bmatrix})$$

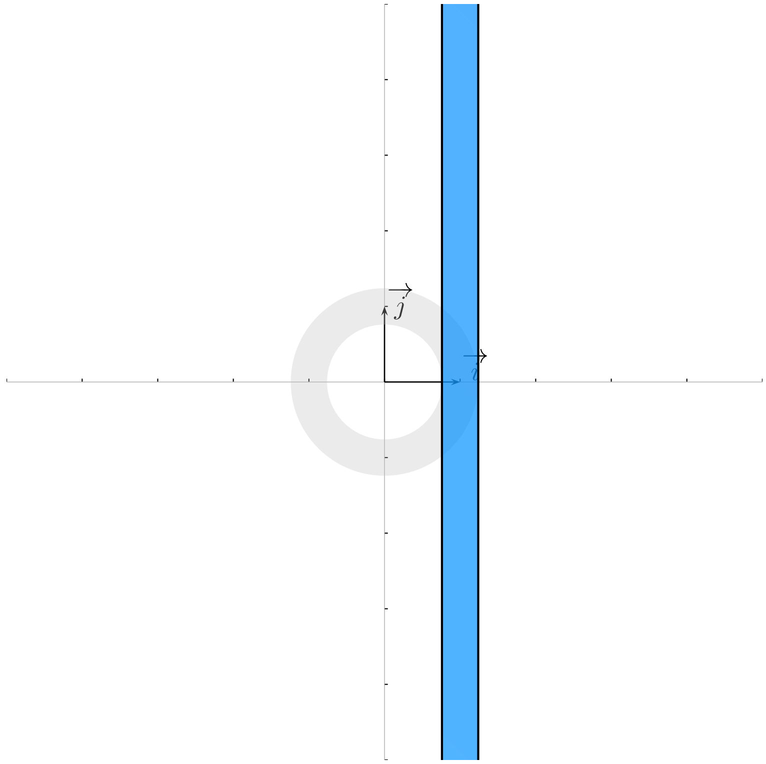


FIGURE 4. 0.95 uncertainty area for the random variable $\begin{bmatrix} x' \\ y' \end{bmatrix}$ after linearization of the system about $\theta = 0$, $\sigma_2^2 = 10000$

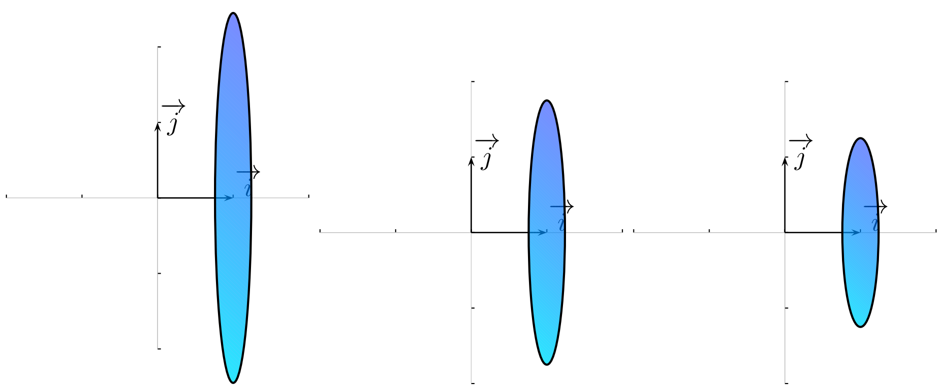


FIGURE 5. 0.95 uncertainty area for the random variable $\begin{bmatrix} x' \\ y' \end{bmatrix}$ after linearization of the system about $\theta = 0$, and using respectively $\sigma_2^2 = 1$,$\sigma_2^2 = 0.5$ and $\sigma_2^2 = 0.25$

In figure @@, Matlab simulation of 1000 realisation of random variable $\begin{bmatrix} x + \cos\theta \\ y + \sin\theta \end{bmatrix}$ are superposed on relevant uncertainty area.

5

Handling an additive constant D_1 in motion model is straightforward.

$$x_t = Ax_{t-1} + \underbrace{Bu_t + D_1}_{\text{deterministic constant}} + \epsilon_t$$

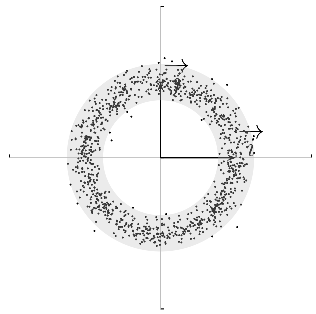


FIGURE 6. Matlab simulation of the random variable $\begin{bmatrix} x' \\ y' \end{bmatrix}$, with $\sigma_2^2 = 10000$

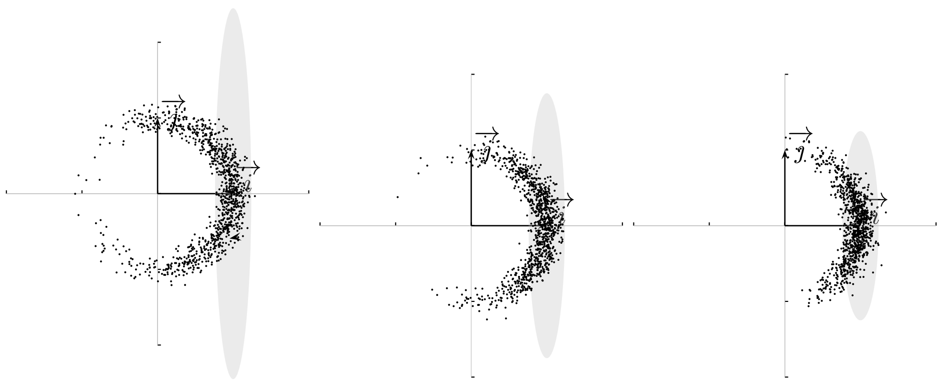


FIGURE 7. Matlab simulation of the random variable $\begin{bmatrix} x' \\ y' \end{bmatrix}$, with $\sigma_2^2 = 1$, $\sigma_2^2 = 0.5$ and $\sigma_2^2 = 0.25$

Since Bu_t is already deterministic in our model, it suffices to replace in the derivation of the motion step Bu_t by $Bu_t + D_1$ to get

$$\bar{\mu}_t = Ax_{t-1} + Bu_t + D_1 + \epsilon_t$$
$$\bar{\Sigma}_t = A\Sigma_{t-1}A^T + R$$

If the measurement model is now

$$z_t = Cx_t + D_2 + \xi_t$$

Suppose we observe instead $z'_t = z_t - D_2$, that is we systematically substract D_2 to actual measurement. We can use the original derivation to get the moments of the posterior:

$$\mu_t = \bar{\mu}_t + K_t(z'_t - C\bar{\mu}_t)$$
$$\Sigma_t = (I - K_tC)\bar{\mu}_t$$

Now replace z'_t by actual measurement.

$$\mu_t = \bar{\mu}_t + K_t(z_t - D_2 - C\bar{\mu}_t)$$
$$\Sigma_t = (I - K_tC)\bar{\mu}_t$$

6

Not sure to understand what is asked exactly in this exercise.

Appendix A. uncertainty ellipse

Appendix B. moments of $\cos\theta$ and $\sin\theta$

Appendix C. image of a disk by $f : (z, \theta) \mapsto z + e^{i\theta}$