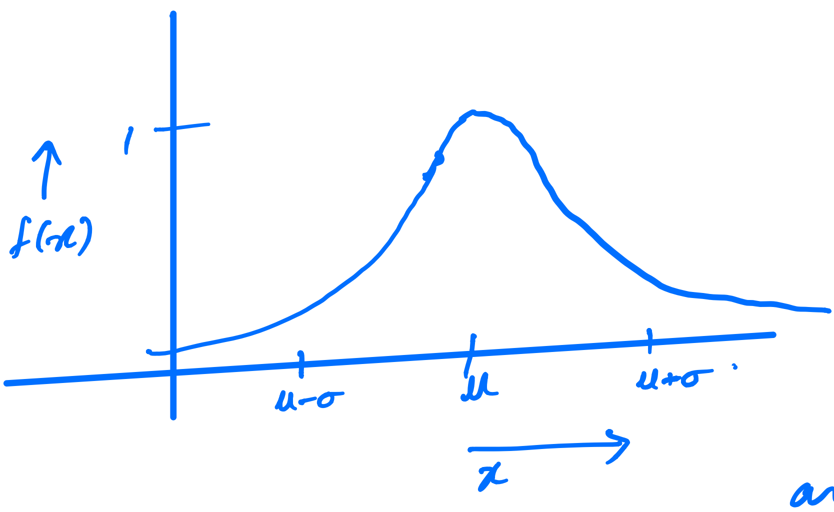
Single Variable Gaussian Probability Density Function:

Consider the function

$$f(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{or} \quad \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$



We call this as a Gaussian function in the variable x. But if x represent a random variable and $f(x)$ represent the chance that the rv takes the value x. then one can call it as a chance function. The chance or the odds that x takes a certain value x_0 or x_1

→ Moreover if we impose $\int_{-\infty}^{\infty} \eta f(x)\, dx = 1.$ and evaluate η to be $\frac{1}{\sqrt{2\pi\sigma^2}}$ then we replace $f(x)$ by $p(x)$ to call it as the probability density function.

→ $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(x-\mu)^2/(2\sigma^2)) \longrightarrow (1)$

is a gaussian pdf in the gaussian rv x

→ Also represented as $N(\mu, \sigma^2)$ is the only distribution to be completly characterized by the mean and variance of the distributon

→ Consider $y = ax$ and let $f(x) = \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$

then $f(x) = \exp\left(-\frac{(y/a - \mu)^2}{2\sigma^2}\right) \longrightarrow (2)$

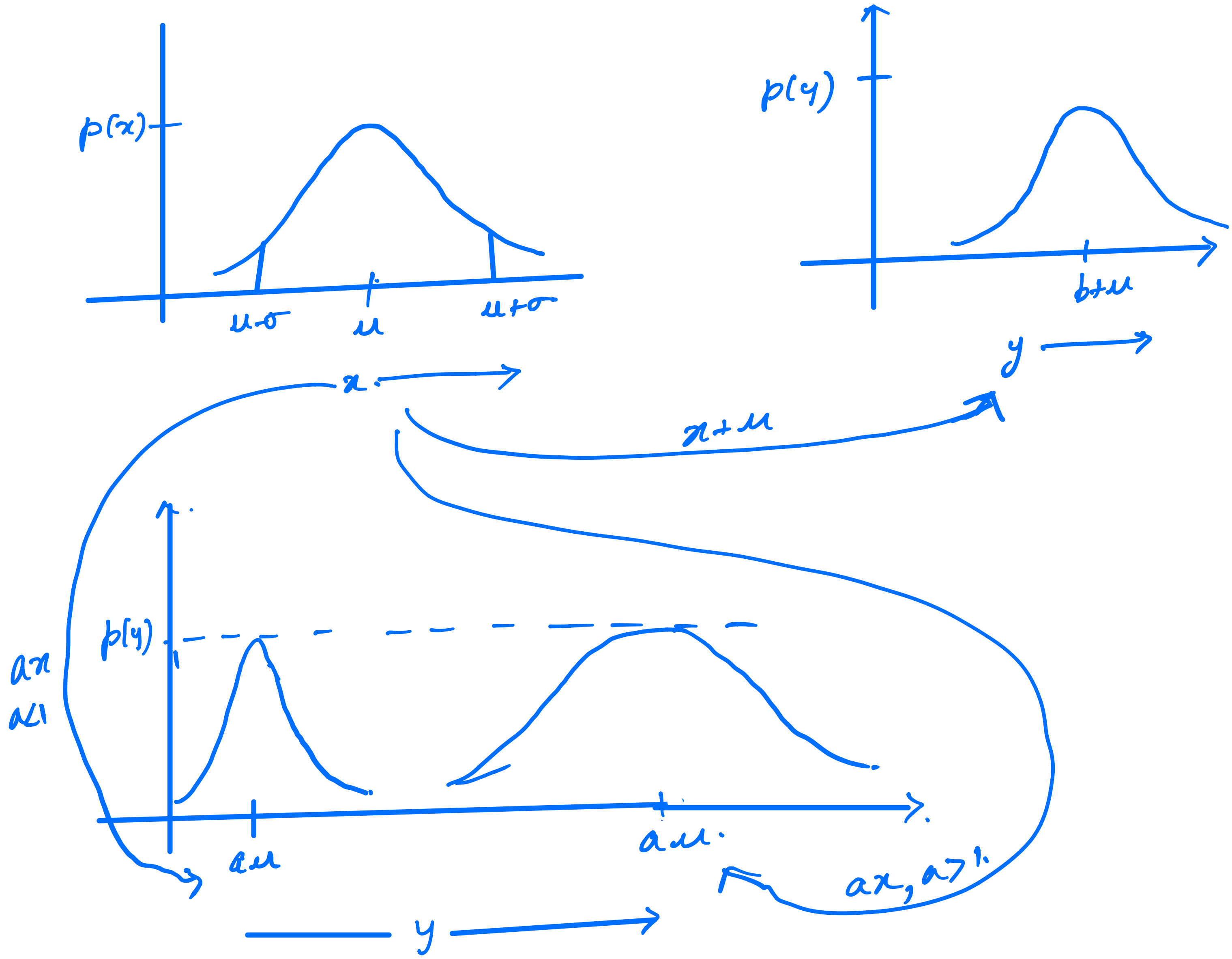
$= \exp\left(-\frac{(y - a\mu)^2}{2\sigma^2 a^2}\right) = g(y). \longrightarrow (3)$

If g where to be a pdf in rv y with some n_1 for which $\int_{-\infty}^{\infty} n_1 g(y) dy = 1$, we get

$n_1 = \frac{1}{\sqrt{2\pi\sigma^2 a^2}} \longrightarrow (4)$

if $x \in N(\mu, \sigma^2)$ and $y = ax$ then $y \in N(a\mu, a^2\sigma^2)$

Similarly if $y = x + b$ then $y \in N(b+\mu, \sigma^2)$.



→ Let $y = ax + b$ or $x = \dfrac{y-b}{a}$

$$f(x) = \exp\left(-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}\right)$$

$$= \exp\left(-\frac{\left(y-(b+a\mu)\right)^2}{2\sigma^2 a^2}\right) = g(y)$$

→ (5)

With appropriate normalization constants then if $x \in N(\mu, \sigma^2)$ and $y = ax + b$ then $y \in N(a\mu + b, a^2\sigma^2)$.

Suppose $y = \cos x$ or $x = \cos^{-1} y$

then $g(y) = \exp\left(\dfrac{-(\cos^{-1} y - \mu)^2}{2\sigma^2}\right) \rightarrow (b)$

Then $g(y)$ is NO MORE a Gaussian function and the associated pdf $p(y)$ is NO MORE a Gaussian pdf.

However $y = \cos x_0 - \sin x_0 (x - x_0)$ is a Taylor series about x_0

or $y = -(\sin x_0)x + (x_0 \sin x_0 + \cos x_0)$

$= ax + b$

Then $y \approx N(a\mu + b, a^2\sigma^2)$.

Multivariate Gaussian Distribution

$$X = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} x \\ y \\ \theta \end{bmatrix}$$

$$p(X) = p(x=x, y=y) = \underbrace{p(x=x) \cdot p(y=y)}_{\rightarrow (7)}$$

(under independence assumption)

$$p(X) = \eta \exp\left[-\left(\underbrace{[x-\mu_x, y-\mu_y]}_{\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}} \underbrace{\begin{bmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{bmatrix}}_{\Sigma^{-1}} \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix} \right) \right] \rightarrow (8).$$

(Σ is the covariance matrix)

$$\eta = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \rightarrow (9)$$

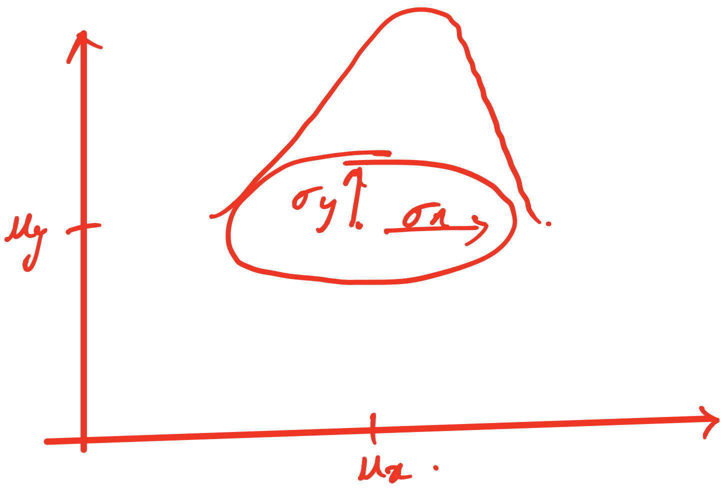
Then we say $X \in N(\mu, \Sigma)$.

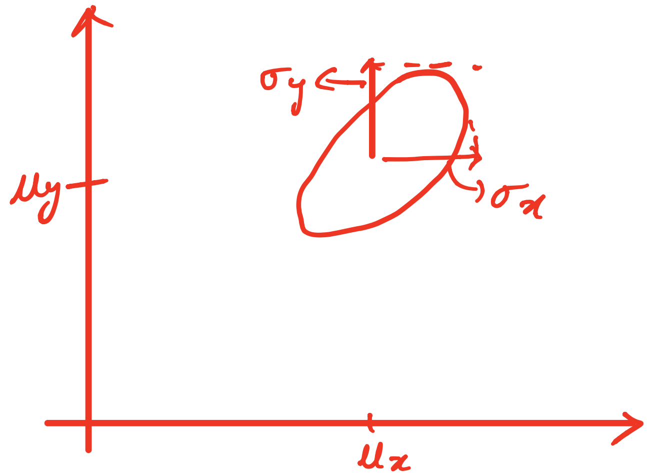
If x and y are corelated then the covariance matrix is $\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y \\ \sigma_y\sigma_x & \sigma_y^2 \end{bmatrix}$ → covariance term.

Then $p(x) = p(x=x, y=y) =$

$$\eta \exp(-[x-\mu]^T \Sigma^{-1} [x-\mu]).$$

where $\mu = [\mu_x \ \mu_y \ \mu_\theta]^T$ or $[\mu_x \ \mu_y]^T$.





Like before if $X_{n\times 1} \approx N(\mu_{n\times 1}, \Sigma_{n\times n})$ and.

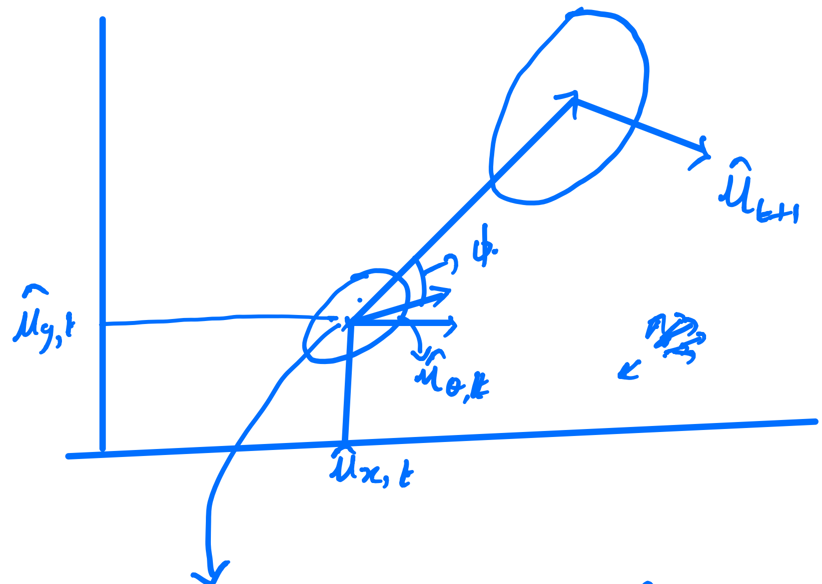
$\underset{n\times 1}{Y} = \underset{n\times n}{A} X + b_{n\times 1}$ then $\boxed{Y \approx N(A\mu + b, A\Sigma A^T)}$.

And if $Y = f(x)$, then

$\boxed{Y \approx N(F\mu + b, F\Sigma F^T)}$, where

$$\boxed{F = \frac{\partial f}{\partial x}\Big|_{x=x_0} \text{ is the Jacobian of } f.}$$

EKF localization / state estimation of a Mobile Robot (Obstacle / object)



$\hat{\mu}_{t+1} \leftarrow [\hat{\mu}_{x\,t+1}, \hat{\mu}_{y\,t+1}, \hat{\mu}_{\theta\,t+1}]$.

The state of the robot X_t at t is a rv, $X_t \in N(\mu_t, \Sigma_t)$

$\hat{\mu}_t \leftarrow [\hat{\mu}_{xt} \ \hat{\mu}_{yt} \ \hat{\mu}_{\theta t}]$.

$$\Sigma_t = \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y & \sigma_x\sigma_\theta \\ \sigma_y\sigma_x & \sigma_y^2 & \sigma_y\sigma_\theta \\ \sigma_\theta\sigma_x & \sigma_\theta\sigma_y & \sigma_\theta^2 \end{bmatrix}$$

state noise

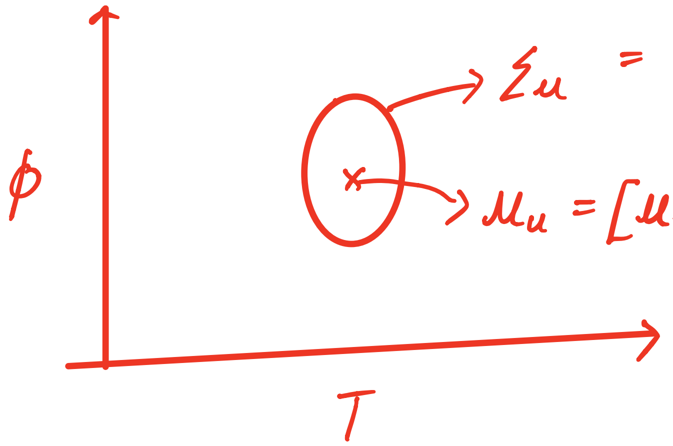
$$P(X_t = \alpha_t) = \eta \exp[\alpha_{xt} - \mu_x \ \ \alpha_{yt} - \mu_y] \Sigma^{-1} [\alpha_{xt} - \mu_x \ \ \alpha_{yt} - \mu_y]^T$$

Now if the robot at time t is subject to a control $u = [T \ \phi]^T$ s.t

$$\hat{X}_{t+1} = f(x_t, u_{t+1}) =$$

$$\begin{bmatrix} \hat{x}_{t+1} \\ \hat{y}_{t+1} \\ \hat{\theta}_{t+1} \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \\ \theta_t \end{bmatrix} + \begin{bmatrix} T\cos(\theta_t + \phi) \\ T\sin(\theta_t + \phi) \\ \phi \end{bmatrix} \rightarrow (1)$$

What is the characterization of pdf of $\hat{X}_{t+1}$ if $u \in N(\mu_u, \Sigma_u)$



$\Sigma_u = \begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix} \rightarrow$ control noise

$\mu_u = [\mu_T \ \mu_\phi]^T$

(i) If x_t evolves to $\hat{x}_{t+1}$ due to a $u_{t+1} \in N(\mu_u, \Sigma_u)$

then μ_t evolves to a $\hat{\mu}_{t+1}$ due to a $\mu_{u_{t+1}}$

as

$$\hat{\mu}_{t+1} = \begin{bmatrix} \hat{\mu}_{x,t+1} \\ \hat{\mu}_{y,t+1} \\ \hat{\mu}_{\theta t+1} \end{bmatrix} = \mu_{t_{3\times1}} + \begin{bmatrix} T\cos(\mu_{\theta t} + \phi) \\ T\sin(\mu_{\theta t} + \phi) \\ \phi \end{bmatrix} \quad (2)$$

The above is of the form $\hat{\mu}_{t+1} = f(\mu_t, \mu_{ut})$

$\Sigma_{u,t}$ is the control noise

$$F = \frac{\partial f}{\partial \mu_t} = \begin{bmatrix} \frac{\partial \hat{\mu}_{x,t+1}}{\partial \mu_{xt}} & \frac{\partial \hat{\mu}_{xt+1}}{\partial \mu_{yt}} & \frac{\partial \hat{\mu}_{\theta t+1}}{\partial \mu_{\theta t}} \\ \frac{\partial \hat{\mu}_{yt+1}}{\partial \mu_{xt}} & \cdots & \frac{\partial \hat{\mu}_{yt+1}}{\partial \mu_{\theta t}} \\ \frac{\partial \hat{\mu}_{\theta t+1}}{\partial \mu_{xt}} & \cdots & \frac{\partial \hat{\mu}_{\theta t+1}}{\partial \mu_{\theta t}} \end{bmatrix} \quad (3)$$

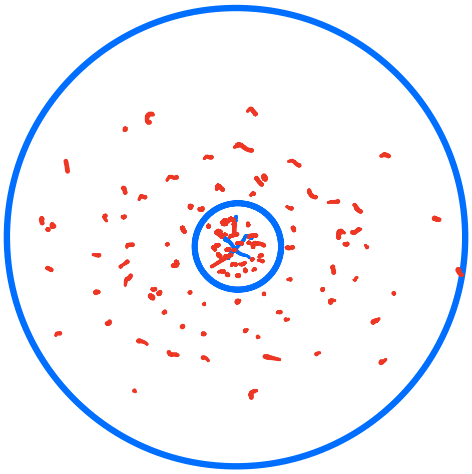
$$= \begin{bmatrix} 1 & 0 & -T\sin(\mu_{\theta t} + \phi) \\ 0 & 1 & T\cos(\mu_{\theta t} + \phi) \\ 0 & 0 & 1 \end{bmatrix} \longrightarrow (4)$$

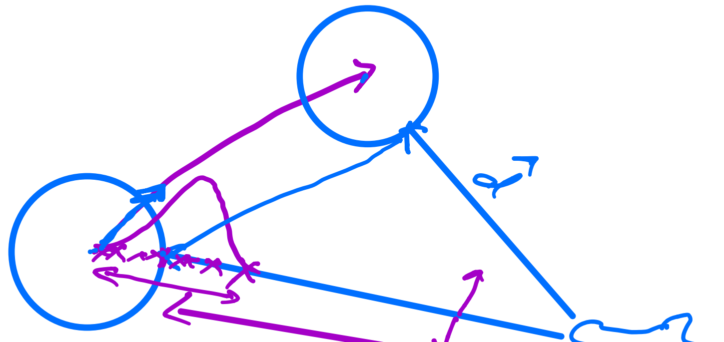
$$G = \frac{\partial f}{\partial \mu_{u\,t+1}} = \begin{bmatrix} \cos(\mu_{\theta t} + \phi) & -T\sin(\mu_{\theta t} + \phi) \\ \sin(\mu_{\theta t} + \phi) & T\cos(\mu_{\theta_t} + \phi) \\ 0 & 1 \end{bmatrix} \longrightarrow (6).$$

$$\begin{bmatrix} \frac{\partial \hat{\mu}_{x\,t+1}}{\partial T} & \frac{\partial \hat{\mu}_{x\,t+1}}{\partial \phi} \\ \vdots & \\ \frac{\partial \hat{\mu}_{\theta\,t+1}}{\partial T} & \frac{\partial \hat{\mu}_{\theta\,t+1}}{\partial \phi} \end{bmatrix} \longrightarrow (5)$$

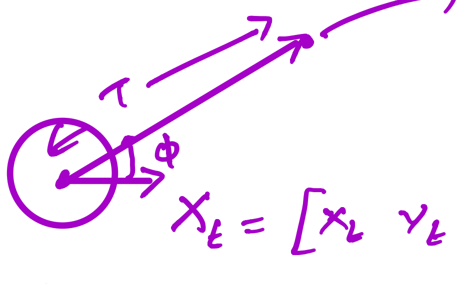
Then
$$\underset{[3,3]}{\hat{\Sigma}_{t+1}} = \underset{[3,3]}{F}\ \underset{[3,3]}{\Sigma_t}\ \underset{[3,3]}{F^T} + \underset{[3,2]}{G}\ \underset{[2,2]}{\Sigma_{u_{t+1}}}\ \underset{[2,3]}{G^T}. \longrightarrow (6)$$

$$\begin{bmatrix} \sigma_{x_{t+1}}^2 & \sigma_x\sigma_y & \sigma_x\sigma_\theta \\ \sigma_y\sigma_x & \sigma_{y_{t+1}}^2 & \sigma_y\sigma_\theta \\ \sigma_\theta\sigma_x & \sigma_\theta\sigma_y & \sigma_\theta^2 \end{bmatrix} = \underset{3\times5}{\begin{bmatrix} F & G \end{bmatrix}} \underset{[5,5]}{\begin{bmatrix} \Sigma_t & 0 \\ 0 & \Sigma_{u_{t+1}} \end{bmatrix}} \underset{[5,3]}{\begin{bmatrix} F^T \\ G^T \end{bmatrix}}$$





$d \in N(\mu_d, \sigma_d^2)$



$\to T\cos\phi, T\sin\phi$ or $T\cos(\theta+\phi), T\sin(\theta+\phi)$

$X_t = [x_t \;\; y_t \;\; \theta_t]^T.$

or $\cancel{x_t} + T\cos(\theta+\phi)$

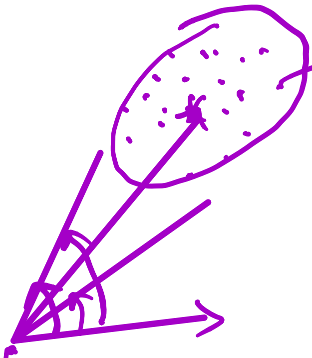
$y_t + T\sin(\theta+\phi)$

in the absence of noise in T, ϕ

$X_{t+1} = f(X_t, u_{t+1}) \quad u_{t+1} = [T \;\; \phi]^T.$

$\hookrightarrow \in N(\mu_{u_{t+1}}, \Sigma_{u_{t+1}})$

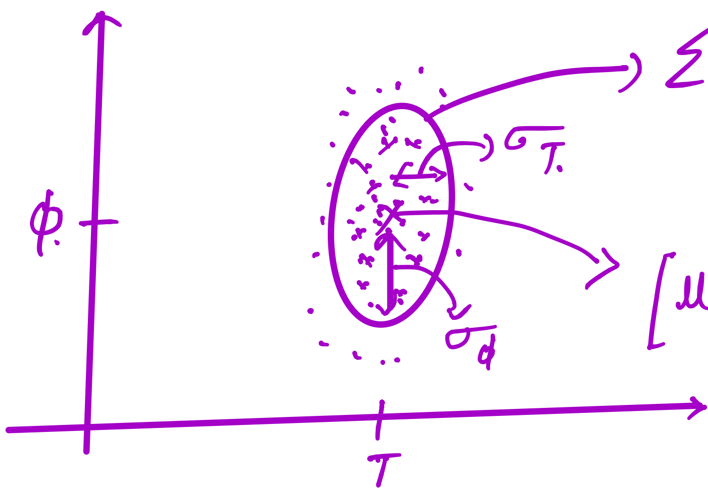
or $N(\mu_u, \Sigma_u).$



how do you characterize the distribution?

even if X_t is deterministically known, we cannot say precisely where X_{t+1} is?

Starting from $u \in N(\mu_u, \Sigma_u)$



$$\Sigma_u = \begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix}$$

$$[\mu_T, \mu_\phi]^T \leftarrow \mu_u$$

$$\hat{X}_{t+1} = f(X_t, u_{t+1}).$$

equivalently

$$\frac{df}{d\mu_{u_{t+1}}}$$

$$\hat{\mu}_{t+1} = f(\mu_t, \mu_{u_{t+1}}).$$

$$\begin{bmatrix} \hat{\mu}_{x\,t+1} \\ \hat{\mu}_{y\,t+1} \\ \hat{\mu}_{\theta\,t+1} \end{bmatrix} = \begin{bmatrix} \mu_{xt} \\ \mu_{yt} \\ \mu_{\theta t} \end{bmatrix} + \begin{bmatrix} T\cos(\mu_{\theta t} + \phi). \\ T\sin(\mu_{\theta t} + \phi). \\ \phi \end{bmatrix}$$

Mean evolution

What is the variance or covariance evolution?

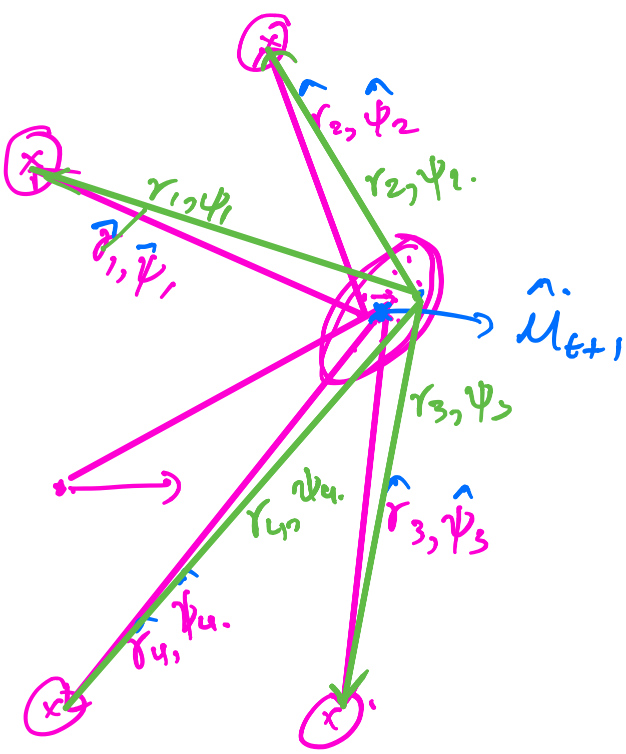
What is $\hat{\Sigma}_{t+1}$?

or what is $\hat{\Sigma}_{t+1}$ in terms of $\Sigma_{u_{t+1}}$?

If $u_{t+1} \in N(\mu_{t+1}, \Sigma_{u_{t+1}})$.

$$\mu_{t+1} \to \begin{bmatrix} T \\ \phi \end{bmatrix} \qquad \Sigma_{u_{t+1}} \to \begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix}$$

μ_t —



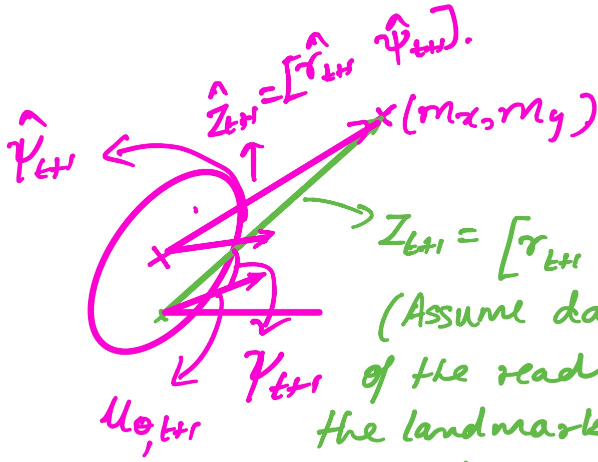
If $\hat{Z}_{t+1} = \begin{bmatrix} \hat{r}_{t+1} \\ \hat{\psi}_{t+1} \end{bmatrix}$

$= h(\hat{\mu}_{t+1})$.

$Z_{t+1} = \begin{bmatrix} r_{t+1} \\ \psi_{t+1} \end{bmatrix}$

$\hat{Z}_{t+1} - Z_{t+1}$

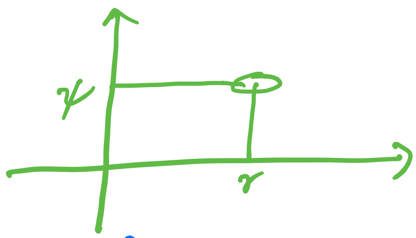
Consider just one landmark i present.



$Z_{t+1} = [r_{t+1} \ \psi_{t+1}]$.
(Assume data assoc of the reading Z_{t+1} to the landmark l_i at (m_x, m_y) is given by some algo)

$$\hat{Z}_{t+1} = \begin{bmatrix} \hat{r}_{t+1} \\ \hat{\psi}_{t+1} \end{bmatrix} = \begin{bmatrix} \sqrt{(m_x - \hat{u}_{x,t+1})^2 + (m_y - \hat{u}_{y,t+1})^2} \\ \tan^{-1}\left(\frac{m_y - \hat{u}_{y,t+1}}{m_x - \hat{u}_{x,t+1}}\right) - \hat{u}_{\theta,t+1} \end{bmatrix} \longrightarrow (7).$$

$$Q = \text{measurement noise} = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\psi^2 \end{bmatrix}$$



We can write $\hat{Z}_{t+1} = h(\hat{u}_{t+1})$. $\longrightarrow (8)$ as seen in (7)

$H \rightarrow$ measurement Jacobian $= \dfrac{dh}{d\hat{u}_{t+1}}$

$$= \begin{bmatrix} \dfrac{\partial \hat{r}_{t+1}}{\partial u_{x\,t+1}} & \dfrac{\partial \hat{r}_{t+1}}{\partial u_{y\,t+1}} & \dfrac{\partial \hat{r}_{t+1}}{\partial u_{\theta\,t+1}} \\ \dfrac{\partial \hat{\psi}_{t+1}}{\partial u_{x\,t+1}} & \dfrac{\partial \hat{\psi}_{t+1}}{\partial u_{y\,t+1}} & \dfrac{\partial \hat{\psi}_{t+1}}{\partial u_{\theta\,t+1}} \end{bmatrix} \longrightarrow (9)$$

Then $S \rightarrow$ innovation covariance that transforms state covariance to the space of measurement

as $$\underset{(2\times2)}{S} = \underset{(2\times3)}{H}\ \underset{(3\times3)}{\hat{\Sigma}_{t+1}}\ \underset{(3\times2)}{H^T} + \underset{(2\times2)}{Q} \longrightarrow (10).$$

In general for n landmarks we have

$Q = (2n \times 2n)$. $H = (2n \times 3)$, $\hat{Z} = (2n \times 1)$.

Kalman gain (obtained from minimum variance estimate) $K = \hat{\Sigma}_{t+1} H^T S^{-1}$. $\longrightarrow (11)$.

$(3\times3) \times (3\times2n) \times (2n\times2n) = (3\times2n)$

Update equation:

$$\Sigma_{t+1} = \hat{\Sigma}_{t+1} [I - KH] \longrightarrow (12). \quad \text{(covariance update)}$$

(3×3) (3×3) $(3\times2n)\times(2n\times2n)$.

$$\mu_{t+1} = \hat{\mu}_{t+1} + K(Z_{t+1} - \hat{Z}_{t+1}) \longrightarrow (13) \quad \text{(mean update)}$$

→ The update results in a better estimate of the true mean or the actual state of the robot.

→ In a reduced uncertainty

What is really happening here?

Consider $S = H\Sigma H^T + Q$. When the measurement noise Q is much less than state noise or $\underline{Q << H\hat{\Sigma}H^T}$, then $K = \hat{\Sigma} H^T S^{-1}$

$$= \hat{\Sigma} H^T (H\Sigma H^T + Q)^{-1}$$

$$= \hat{\Sigma} H^T (H\Sigma H^T)^{-1} = \hat{\Sigma} H^T (H^T)^{-1} \Sigma^{-1} H^{-1}$$

$$= \underline{\underline{H^{-1}}} \longrightarrow (14)$$

Then $u = \hat{u} + K(z - \hat{z})$ becomes

$$\hat{u} + H^{-1}(z - \hat{z})$$

$$= \cancel{\hat{u}} + H^{-1}z - \cancel{H^{-1}\hat{z}}$$

$$= H^{-1}z \longrightarrow (15)$$

($\because$ $\hat{z} = H\hat{u}$ under 1st order approximation as $\hat{z} = h(\hat{u})$ can be approximated as.

$\hat{z} = H\hat{u}$)

Then we trust the measurement completely as $\underline{\underline{u = H^{-1}z}}$, would mean we have completely dismissed $\hat{u}$ the prediction as the covariance values are high or the state has a flat distribution.

What happens if $H \Sigma H^T << Q$. (state noise is much less than measurement noise)

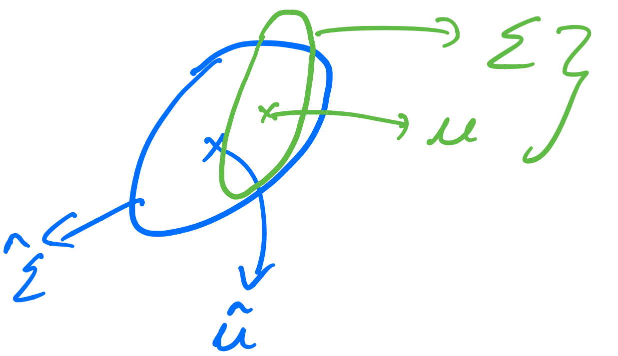
Then $u = \hat{u} + K(z - \hat{z})$.

where $K = \Sigma H^T (H^T \Sigma H + Q)^{-1} = \Sigma H^T Q^{-1} \approx 0$

or $u = \hat{u} \longrightarrow (16)$.

Or we don't trust the measurement at all or we only go by the <u>prediction model</u> and <u>NOT</u> by the <u>measurement model</u>.

Typically after measurement update the uncertainty reduces. The new μ or the new mean is a combination of the prediction $\hat{\mu}$ and the innovation $z-\hat{z}$, weighed by the Kalman gain K.



Simplified Understanding: Scalar Kalman Filter or Scalar EKF.

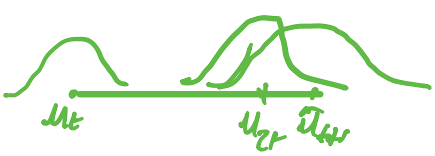
Scalar KF:

Prediction:

$$\hat{\mu}_{t+1} = \mu_t + u$$

$$\hat{\sigma}_{t+1} = \sigma_t^2 + \sigma_u^2$$

Measurement:



$z_t = \mu_{z_t}$ (actual)

$\hat{\mu}_{z_t}$ (predicted)

σ_z^2 (measurement variance)

Multivariate KF / EKF

$$\hat{\mu}_{t+1} = F\mu_t + G u_{t+1}$$

$$\hat{\Sigma}_{t+1} = F\Sigma F^T + G\Sigma_u G^T$$

$\hat{z}_t = H\hat{\mu}_t$ (prediction)

z_t (actual measurement)

Q. . .

$$K = \frac{\hat{\sigma}_{t+1}^2}{\hat{\sigma}_{t+1}^2 + \sigma_z^2}$$

$$K = \Sigma H^T S^{-1}.$$

$$(S = H \Sigma H^T + Q).$$

Update eqns:

Mean update:

$$\mu_{t+1} = \hat{\mu}_{t+1} + K(\mu_{z_{t+1}} - \hat{\mu}_{z_{t+1}})$$

$$\mu_{t+1} = \hat{\mu}_{t+1} + K(\dot{Z}_{t+1} - \hat{Z}_{t+1}).$$

$$\sigma_{t+1}^2 = \hat{\sigma}_{t+1}^2 (1 - K).$$

$$\Sigma_{t+1} = \hat{\Sigma}_{t+1} [I - KH].$$