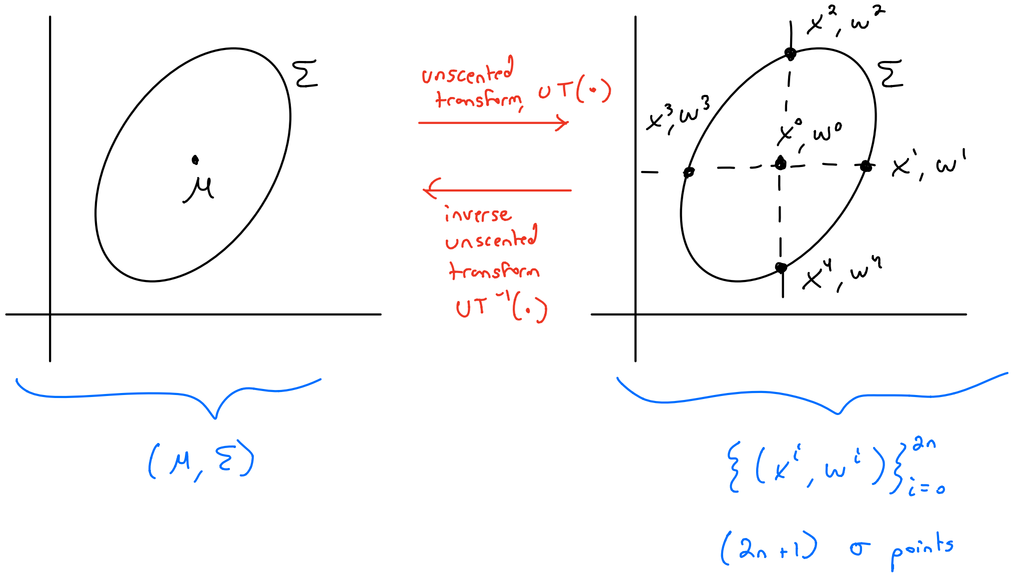
AA 273 Lecture 12 Notes

Tamas Kis

Unscented (σ-point) Transform:



free parameter, often $\lambda = 2$

$$UT(\mu, \Sigma) = \begin{cases} x^0 = \mu, \quad w^0 = \dfrac{\lambda}{\lambda+n} \\ \vdots \\ x^i = \mu + \left(\sqrt{(\lambda+n)\Sigma}\right)_i \\ x^{i+n} = \mu - \left(\sqrt{(\lambda+n)\Sigma}\right)_i \\ w^i = \dfrac{1}{2(\lambda+n)} \end{cases}$$

ith column of matrix square root

$$UT^{-1}\left(\{x_i, w_i\}_{i=0}^{2n}\right) = \begin{cases} \mu = \sum_{i=0}^{2n} w^i x^i \\ \Sigma = \sum_{i=0}^{2n} w^i (x^i - \mu)(x^i - \mu)^T \end{cases}$$

empirical mean and covariance of σ-points

Let's prove $UT^{-1}(UT(\mu, \Sigma)) = (\mu, \Sigma)$.

Mean:

$$\sum_{i=0}^{2n} w^i x^i = \frac{\lambda}{\lambda+n}\mu + \sum_{i=1}^{n} \frac{1}{2(\lambda+n)}\left(\mu + \cancel{\left(\sqrt{(\lambda+n)\Sigma}\right)_i}\right) + \sum_{i=n+1}^{2n} \frac{1}{2(\lambda+n)}\left(\mu - \cancel{\left(\sqrt{(\lambda+n)\Sigma}\right)_i}\right)$$

$$= \left(\frac{\lambda}{\lambda+n}\right)\mu + \left[\frac{n}{2(\lambda+n)}\right]\mu + \left[\frac{n}{2(\lambda+n)}\right]\mu = \left[\frac{\lambda}{\lambda+n} + \frac{n}{\lambda+n}\right]\mu = \left(\frac{\lambda+n}{\lambda+n}\right)\mu = \mu$$

Covariance:

$$\sum_{i=0}^{2n} w^i (x^i - \mu)(x^i - \mu)^T = \left(\frac{\lambda}{\lambda+n}\right)\cancel{(\mu - \mu)}^{0}(\mu-\mu)^T + \sum_{i=1}^{n} \frac{1}{2(\lambda+n)} \left(\sqrt{(\lambda+n)\Sigma}\right)_i \left(\sqrt{(\lambda+n)\Sigma}\right)_i^T$$

$$+ \sum_{i=n+1}^{2n} \frac{1}{2(\lambda+n)} \left(-\sqrt{(\lambda+n)\Sigma}\right)_i \left(-\sqrt{(\lambda+n)\Sigma}\right)_i^T$$

$$= \sum_{i=1}^{n} \frac{1}{\lambda+n} \left(\sqrt{(\lambda+n)\Sigma}\right)_i \left(\sqrt{(\lambda+n)\Sigma}\right)_i^T \quad \left\{ \sqrt{\lambda+n} = \text{constant} \right.$$

$$= \sum_{i=1}^{n} (\sqrt{\Sigma})_i (\sqrt{\Sigma})_i^T \overset{?}{=} \Sigma$$

Recall from linear algebra, for any matrix M

$$M = \begin{bmatrix} | & & | \\ m_1 & \cdots & m_n \\ | & & | \end{bmatrix}$$

$$MM^T = \begin{bmatrix} | & & | \\ m_1 & \cdots & m_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & m_1^T & — \\ & \vdots & \\ — & m_n^T & — \end{bmatrix}$$

$$= \sum_{i=1}^{n} m_i m_i^T$$

For us,

$$\sum_{i=1}^{n} (\sqrt{\Sigma})_i (\sqrt{\Sigma})_i^T = (\sqrt{\Sigma})(\sqrt{\Sigma})^T = \Sigma$$

Unscented Kalman Filter:

Predict: UT: $(\mu_{t-1|t-1}, \Sigma_{t-1|t-1}) \xrightarrow{UT} \left\{ (x^i_{t-1|t-1}, w^i) \right\}_{i=0}^{2n}$

σ-point prediction:
$$\left.\begin{aligned} \bar{x}^i_{t|t-1} &= f(x^i_{t-1|t-1}, u_{t-1}) \\ \bar{w}^i &= w^i \end{aligned}\right\} i = 0, \dots, 2n$$

UT^{-1}:
$$\mu_{t|t-1} = \sum_{i=0}^{2n} w^i \bar{x}^i_{t|t-1}$$

$$\Sigma_{t|t-1} = \sum_{i=0}^{2n} w^i (\bar{x}^i_{t|t-1} - \mu_{t|t-1})(\bar{x}^i_{t|t-1} - \mu_{t|t-1})^T + Q_{t-1}$$

Update: UT: $(\mu_{t|t-1}, \Sigma_{t|t-1}) \xrightarrow{UT} \left\{ (x^i_{t|t-1}, w^i) \right\}_{i=0}^{2n}$

σ-point measurement prediction: $y^i_{t|t-1} = g(x^i_{t|t-1}, u_t)$

$$\hat{y}_{t|t-1} = \sum_{i=0}^{2n} w^i y^i_{t|t-1} \quad \Big\}\ \text{expected measurement}$$

Compute Cov(Y) from σ-points: $\Sigma^Y_{t|t-1} = \sum_{i=0}^{2n} w^i \left(y^i_{t|t-1} - \hat{y}_{t|t-1}\right)\left(y^i_{t|t-1} - \hat{y}_{t|t-1}\right)^T + R_t$

(analogous to $C_t \Sigma_{t|t-1} C_t^T + R_t$ in KF)

Compute Cov(X,Y) from σ-points: $\Sigma^{xy}_{t|t-1} = \sum_{i=0}^{2n} w^i \left(x^i_{t|t-1} - \mu_{t|t-1}\right)\left(y^i_{t|t-1} - \hat{y}_{t|t-1}\right)^T$

(analogous to $\Sigma_{t|t-1} C_t^T$ in KF)

Apply Gaussian estimation equations: $\mu_{t|t} = \mu_{t|t-1} + \Sigma^{xy}_{t|t-1}\left(\Sigma^{y}_{t|t-1}\right)^{-1}\left(y_t - \hat{y}_{t|t-1}\right)$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma^{xy}_{t|t-1}\left(\Sigma^{y}_{t|t-1}\right)^{-1}\left(\Sigma^{xy}_{t|t-1}\right)^T$$

(i) λ is a free parameter, typically $\boxed{\lambda = 2}$

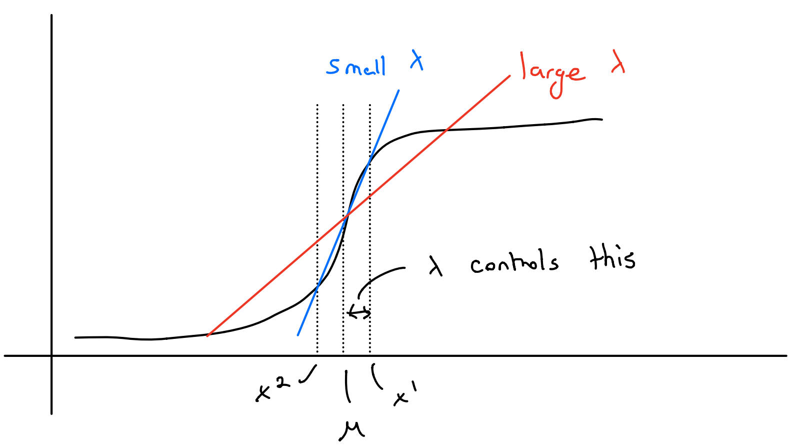
→ UKF essentially performs a finite difference approximation to Jacobians

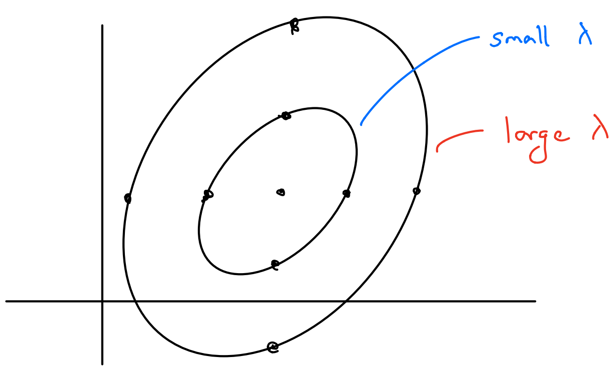
→ λ controls the width of the difference

→ $\lambda \uparrow$ less sensitive to nonlinearities

→ $\lambda \downarrow$ more sensitive to nonlinearities

Graphically,





(ii) Some texts (such as Probabilistic Robotics) use more complex def. of σ-points

→ $(\lambda, \alpha, \kappa, \beta)$ - free parameters

→ constraint: $\lambda = \alpha^2(n+\kappa) - n$ (λ already redundant)

→ $x^i = \underset{\substack{\uparrow \\ \text{scale constant}}}{\mu \pm \alpha} \left(\sqrt{\underset{\substack{\uparrow \\ \text{like our } \lambda}}{(n+\kappa)}\Sigma}\right)_i$

→ one exception for the weight: $w_i^0 = \left(\frac{\lambda}{\lambda+n}\right) + (1-\alpha^2+\beta)$
(only in covariance calc.)

→ stick with the simpler, original form!

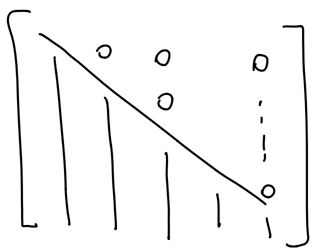
(iii) Which matrix square root to use?

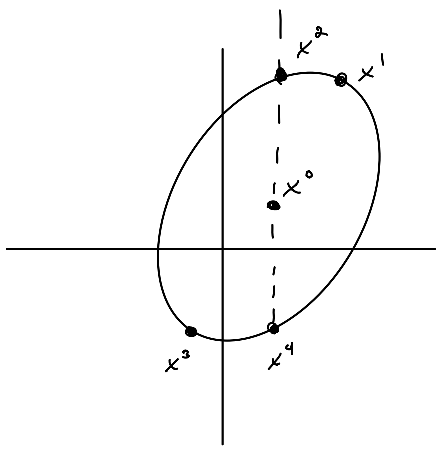
$$x^i = \mu \pm \left(\sqrt{(n+\lambda)\Sigma}\right)_i$$

→ any square root will work

→ Cholesky: $\underset{\substack{\uparrow \\ \text{symmetric P.D.}}}{\Sigma} = \underset{\substack{\uparrow \\ \text{lower triangular}}}{L}L^T$ ← upper triangular

$$\Sigma^{1/2} = L$$

$L =$ 



strange "clumped" distribution of σ-points.

Fast computation: $\frac{1}{6}n^3$

→ SVD square root: $M = \underset{\substack{\uparrow \\ \text{singular value matrix} \\ \to \text{diagonal} \\ \to \text{rectangular}}}{U\Sigma} U^T$ ← orthogonal, square

For us, Σ is symmetric P.D.

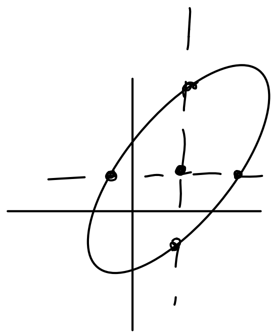
$$\Sigma = U \Lambda U^T$$

U: eigenvector matrix (orthogonal)

Λ: eigenvalue matrix

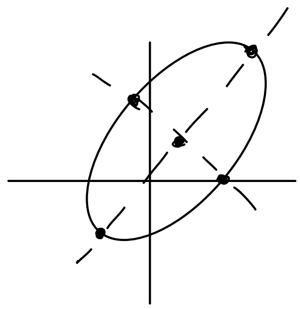
(a) axis-aligned: $\Sigma = (U \Lambda^{1/2} U^T)(U \Lambda^{1/2} U^T)$

$$\Sigma^{1/2} = U \Lambda^{1/2} U^T$$



(b) ellipse-aligned: $\Sigma = (U \Lambda^{1/2})(\Lambda^{1/2} U^T)$

$$\Sigma^{1/2} = U \Lambda^{1/2}$$



more interpretable

better spread

Computation: $4n^3$

(24 x slower than Cholesky)