Approximating statistics with sigma points

- We now look at a different approach to characterizing mean, covariance of output of a nonlinear function
- We avoid Taylor-series expansion; instead, a number of function evaluations are performed, and results are used to compute estimated mean, covariance
- This has several advantages:
  1. Derivatives do not need to be computed (which is one of the most error-prone steps when implementing EKF), also implying
  2. The original functions do not need to be differentiable, and
  3. Better covariance approximations are usually achieved relative to EKF, allowing for better state estimation,
  4. All with comparable computational complexity to EKF

General idea of sigma-point methods

- Steps for characterizing statistics of nonlinear function
  1. Set of sigma points $\mathcal{X}$ is chosen so that (possibly weighted) mean, covariance of points exactly matches mean $\bar{x}$ and covariance $\Sigma_{\tilde{x}}$ of *a priori* RV being modeled (input to function)
  2. These points are then passed through the nonlinear function, resulting in a transformed set of sigma points $\mathcal{Y}$
  3. The *a posteriori* mean $\bar{y}$ and covariance $\Sigma_{\tilde{y}}$ are then approximated by statistical average and covariance of transformed points $\mathcal{Y}$
- Note that the sigma points comprise a fixed small number of vectors that are calculated deterministically—not like particle filter methods

Specific details on creating sigma points

- If input RV $x \in \mathbb{R}^L$ and $x \sim \mathcal{N}(\bar{x}, \Sigma_{\tilde{x}})$, then $p + 1 = 2L + 1$ sigma points are generated as the set (indexed from 0 to p)

$$\mathcal{X} = \left\{\bar{x}, \bar{x} + \gamma\sqrt{\Sigma_{\tilde{x}}}, \bar{x} - \gamma\sqrt{\Sigma_{\tilde{x}}}\right\},$$

  where the matrix square root $R = \sqrt{\Sigma}$ computes a result such that $\Sigma = RR^T$
  - Usually, efficient *Cholesky decomposition* is used, resulting in lower-triangular R

  ⚠ Default in MATLAB/Octave is upper-triangular matrix that must be transposed!
- Weighted mean, covariance of $\mathcal{X}$ equal to original for some $\{\gamma, \alpha^{(\mathrm{m})}, \alpha^{(\mathrm{c})}\}$ if

$$\bar{x} = \sum_{i=0}^{p} \alpha_i^{(\mathrm{m})} \mathcal{X}_i \qquad \text{and} \qquad \Sigma_{\tilde{x}} = \sum_{i=0}^{p} \alpha_i^{(\mathrm{c})} (\mathcal{X}_i - \bar{x})(\mathcal{X}_i - \bar{x})^T,$$

  where $\alpha_i^{(\mathrm{m})}$ and $\alpha_i^{(\mathrm{c})}$ are real scalars where $\alpha_i^{(\mathrm{m})}$ and $\alpha_i^{(\mathrm{c})}$ must both sum to one

Weights for different SPKF methods

- Values used by *Unscented Kalman Filter* (UKF) and *Central Difference Kalman Filter* (CDKF):

Method	γ	$\alpha_0^{(\mathrm{m})}$	$\alpha_k^{(\mathrm{m})}$	$\alpha_0^{(\mathrm{c})}$	$\alpha_k^{(\mathrm{c})}$
UKF	$\sqrt{L+\lambda}$	$\frac{\lambda}{L+\lambda}$	$\frac{1}{2(L+\lambda)}$	$\frac{\lambda}{L+\lambda}+(1-\alpha^2+\beta)$	$\frac{1}{2(L+\lambda)}$
CDKF	h	$\frac{h^2-L}{h^2}$	$\frac{1}{2h^2}$	$\frac{h^2-L}{h^2}$	$\frac{1}{2h^2}$

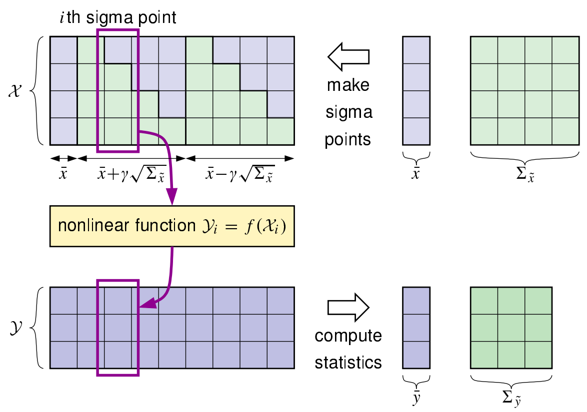
$\lambda = \alpha^2(L+\kappa)-L$, with $(10^{-2} \le \alpha \le 1)$. This α is different from $\alpha^{(\mathrm{m})}$, $\alpha^{(\mathrm{c})}$. $\kappa \in \{0, 3-L\}$

For Gaussian RVs, $\beta = 2$. h may take any positive value; for Gaussian RVs, $h = \sqrt{3}$

- UKF, CDKF derived very differently, but final methods essentially identical
- CDKF has only one "tuning parameter" h, so implementation is simpler; also has higher theoretic accuracy than UKF

Process illustrated

- Diagram illustrates the overall process, with the sets $\mathcal{X}$ and $\mathcal{Y}$ stored compactly with each set member a column in a matrix



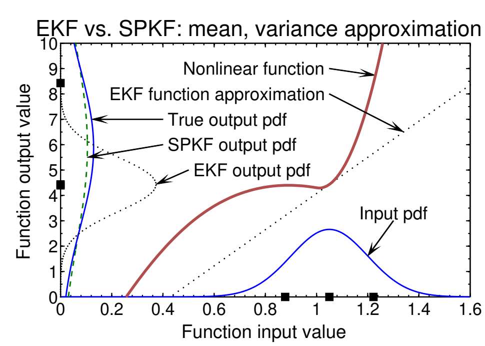
- Create input sigma points
- Compute output sigma points $\mathcal{Y}_i = f(\mathcal{X}_i)$
- Compute output mean, covariance

$$\bar{y} = \sum_{i=0}^{p} \alpha_i^{(\mathrm{m})} \mathcal{Y}_i$$

$$\Sigma_{\tilde{y}} = \sum_{i=0}^{p} \alpha_i^{(\mathrm{c})} (\mathcal{Y}_i - \bar{y})(\mathcal{Y}_i - \bar{y})^T$$

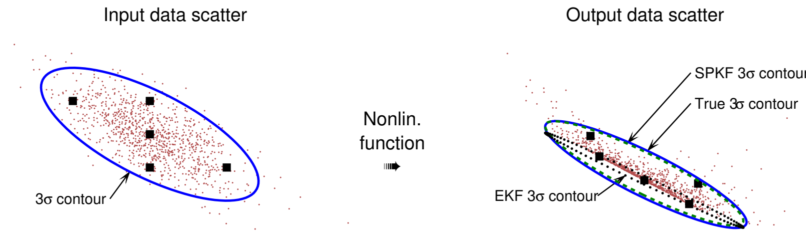
Illustrating improvement in 1-d

- Before introducing SPKF algorithm, we re-examine prior 1-d/ 2-d examples using sigma-point methods
- In 1-d example, $2L+1 = 3$ input sigma points are needed, map to three output sigma points
- Gaussian pdf having mean, variance computed by sigma-point method is shown as a dashed-line PDF
- Closely matches Gaussian pdf having true mean, variance



Illustrating improvement in 2-d

- For 2-d example, $2L + 1 = 5$ sigma points represent input RV
- Similarly, five sigma points represent output RV
- Mean, covar of output sigma points (dashed) closely match true mean, covar



Summary

- EKF approach assumes $\mathbb{E}[\text{fn}(x)] \approx \text{fn}(\mathbb{E}[x])$, and linearizes Taylor-series expansions (analytic linearization)
- Sigma-point approach uses a small number of function evaluations to find mean, covariance (statistical linearization)
- Examples demonstrate that sigma-point approach can provide better mean, covariance estimates than analytic-linearization approach
- Will sigma-point method always be so much better?
  - Answer depends on degree of nonlinearity of the state and output equations—the more nonlinear the better SPKF should be with respect to EKF