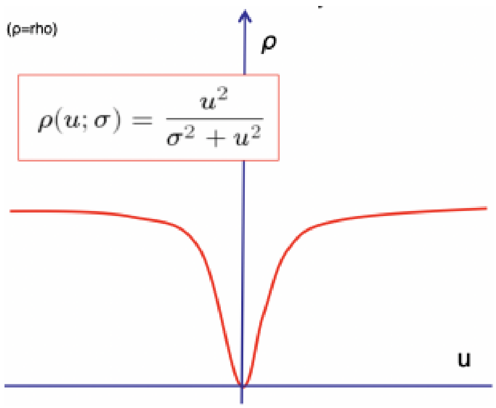
Robust Fitting

In camera calibration or multi-view geometry, we usually need to solve a least-squares problem. In practice, least-squares fitting handles noisy data well but is susceptible to outliers.

Robust Cost Functions

The quadratic growth of the squared error $C(u_i) = u_i^2$ that we've been using so far means that outliers with large residuals u_i exert an outsized influence on cost minimum. We can penalize large residuals (outliers) less by a robust cost function, such as

$$C(u_i, \sigma) = \frac{u_i^2}{\sigma^2 + u_i^2}$$



When the residual u_i is large, the cost C saturates to 1 such that their contribution to the cost is limited, but when u is small, the cost function resembles the squared error.

RANdom SAmple Consensus - RANSAC

RANSAC is an iterative method for estimating the parameters of a mathematical model from a set of observed data containing outliers.

- Robust method (handles up to 50% outliers)
- The estimated model is random but reasonable
- The estimation process divides the observed data into inliers and outliers
- Usually an improved estimate of the model is determined based on the inliers using a less robust estimation method, e.g. least squares

> **Objective**:
> To robustly fit a model $y = f(x, \alpha)$ to a data set S containing outliers
> **Algorithm**:
>
> 1. Estimate the model parameters α_{tmp} from a randomly sampled subset of s data points from S
> 2. Determine the set of inliers $S_{\text{tmp}} \subseteq S$ to be the data points within a distance t of the model
> 3. If this set of inliers is the largest so far, let $S_{\text{IN}} = S_{\text{tmp}}$ and let $\alpha = \alpha_{\text{tmp}}$
> 4. If $|S_{\text{IN}}| < T$, where T is some threshold value, repeat steps 1-3. otherwise stop
> 5. After n trials, stop

Analysis

We can estimate the number of iterations n to guarantee with probability p at least one random sample with an inlier set free of outliers for a given s (minimum number of points required to fit a model) and $\epsilon \in [0,1]$(proportion of inliers)

- The probability that a single random sample contains all inliers is ϵ^s.
- The probability that a single random sample contains at least one outlier is $1-\epsilon^s$.
- The probability that at all n samples contain at least one outlier is $(1-\epsilon^s)^n$.
- The probability that at least one of the n samples does not contain any outliers is $1-(1-\epsilon^s)^n$. Thus

$$p = 1-(1-\epsilon^s)^n$$

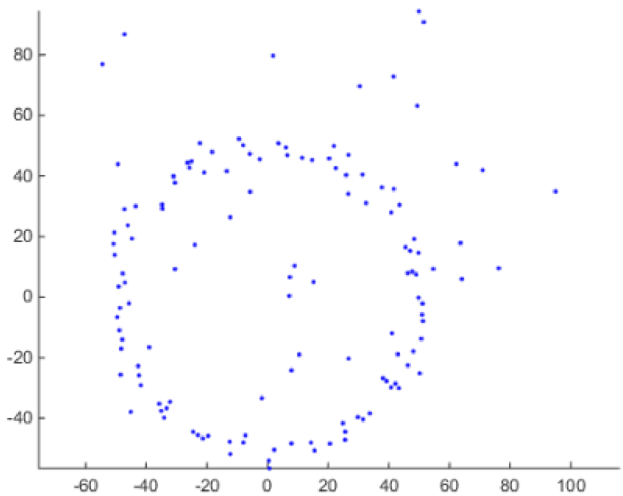
and

$$n = \frac{\log(1-p)}{\log(1-\epsilon^s)}$$

Adaptive RANSAC

> **Objective**:
> To robustly fit a model $y = f(x, \alpha)$ to a data set S containing outliers
> **Algorithm**:
>
> 1. Let $n = \infty$, $S_{\mathrm{IN}} = \emptyset$ and $\#\mathrm{iterations} = 0$.
> 2. While $n > \#\mathrm{iterations}$, repeat 3-5.
> 3. Estimate parameters α_{tmp} from a random s-tuple from S.
> 4. Determine inlier set S_{tmp}, i.e. data points within a distance t of the model $y = f(x, \alpha)$.
> 5. If $|S_{\mathrm{tmp}}| > |S_{\mathrm{IN}}|$, set $S_{\mathrm{IN}} = S_{\mathrm{tmp}}$, $\alpha = \alpha_{\mathrm{tmp}}$, $\epsilon = \frac{|S_{\mathrm{IN}}|}{|S_{\mathrm{tmp}}|}$ and $\frac{\log(1-p)}{\log(1-\epsilon^s)}$ with $p = 0.99$ or higher. Increase $\#\mathrm{iterations}$ by 1.

Example

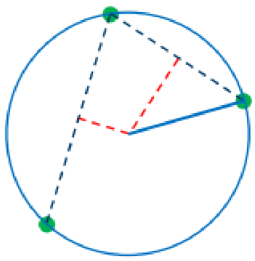


We want to fit a circle $(x-x_0)^2 + (y-y_0)^2 = r^2$ to these data points by estimating the 3 parameters x_0, y_0 and r. The data consists of some points on a circle with Gaussian noise and some random points.

To estimate the circle using RANSAC, we need two things:

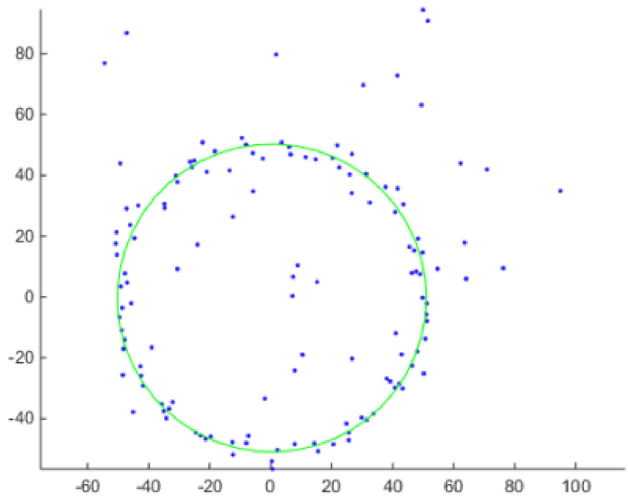
- A way to estimate a circle from s-points, where s is as small as possible.

- The smallest number of points required to determine a circle is 3, i.e. $s = 3$, and the algorithm for computing the circle is quite simple.



- A way to determine which of the points are inliers for an estimated circle.
  - The distance from a point (x_i, y_i) to a circle $(x - x_0)^2 + (y - y_0)^2 = r^2$ is given by $|\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - r|$
  - So for a threshold value t, we say that (x_i, y_i) is an inlier if $|\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - r| < t$

The RANSAC algorithm evaluates many different circles and returns the circle with the largest inlier set



Reference

Robust estimation with RANSAC