Random Sample and Consensus

COSC450

Lecture 4

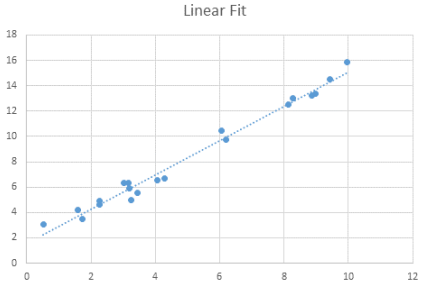
Robust Estimation

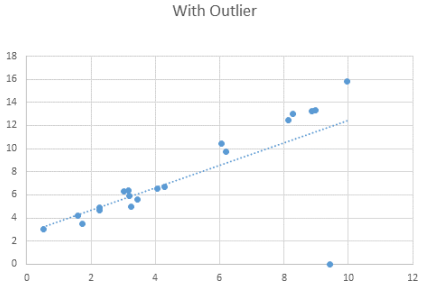
Fitting a model to data points:

- Measurements are uncertain
- Gaussian errors $\Longrightarrow$ least squares fit
- Outliers cause problems

Robust methods:

- Least Median of Squares
- Least Trimmed Squares
- Iteratively Reweighted Least Squares
- RANSAC





RANSAC for Line Fitting

procedure RANSAC(n points, p_i; numTrials; threshold)
    consensus $= \{\}$
    for trial $\leftarrow 1 \ldots$ numTrials **do**
        $s1 =$ rand(1,n) $s2 =$ rand(1,n) thisConsensus $= \{\}$
        $L =$ fitLine(p_{s1}, p_{s2})
        for i $\leftarrow 1 \ldots n$ **do**
            if distance(L, p_i) $<$ threshold **then**
                thisConsensus $=$ thisConsensus $+ p_i$
            end if
        end for
        if $\|$thisConsensus$\| > \|$consensus$\|$ **then**
            consensus $=$ thisConsensus
        end if
    end for
    return fitLine(consensus)
end procedure

General RANSAC

More generally:

- We have a set of n points (or items), p_i
- We can fit a model, M, to $k \ll n$ points
- We can find the distance, $d(M, p)$ between a model and a point
- We also need a threshold for acceptance and a number of trials

RANSAC for Model Fitting

procedure $\mathrm{RANSAC}(n$ points, p_i; numTrials; threshold)
    consensus $= \{\}$
    for trial $\leftarrow 1 \ldots$ numTrials **do**
        $S = \{$ k of the n points chosen at random $\}$
        $M =$ fitModel(S)
        thisConsensus $\leftarrow \{p_i : d(M, p_i) <$ threshold$\}$
        if $\|$thisConsensus$\| > \|$consensus$\|$ **then**
            consensus $=$ thisConsensus
        end if
    end for
    return fitModel(consensus)
end procedure

How Many Trials?

Chances of outliers being a problem:

- Some proportion, ϕ, are inliers
- Probability of k good points is ϕ^k
- Probability of one bad trial is $1 - \phi^k$
- Probability of t bad trials is $(1 - \phi^k)^t$

Pick a small probability, p, that this happens

$$p = (1 - \phi^k)^t$$
$$\log(p) = t \log(1 - \phi^k)$$
$$t = \frac{\log(p)}{\log(1 - \phi^k)}$$

But we don't know ϕ

- Suppose we have n points
- Initialise ϕ to some small value
- If we find a consensus set of size c

$$\phi = \max\left(\phi, \frac{c}{n}\right)$$

As we make more trials

- We find larger consensus sets
- ϕ increases, so t decreases
- When t is less than the number of trials done, stop