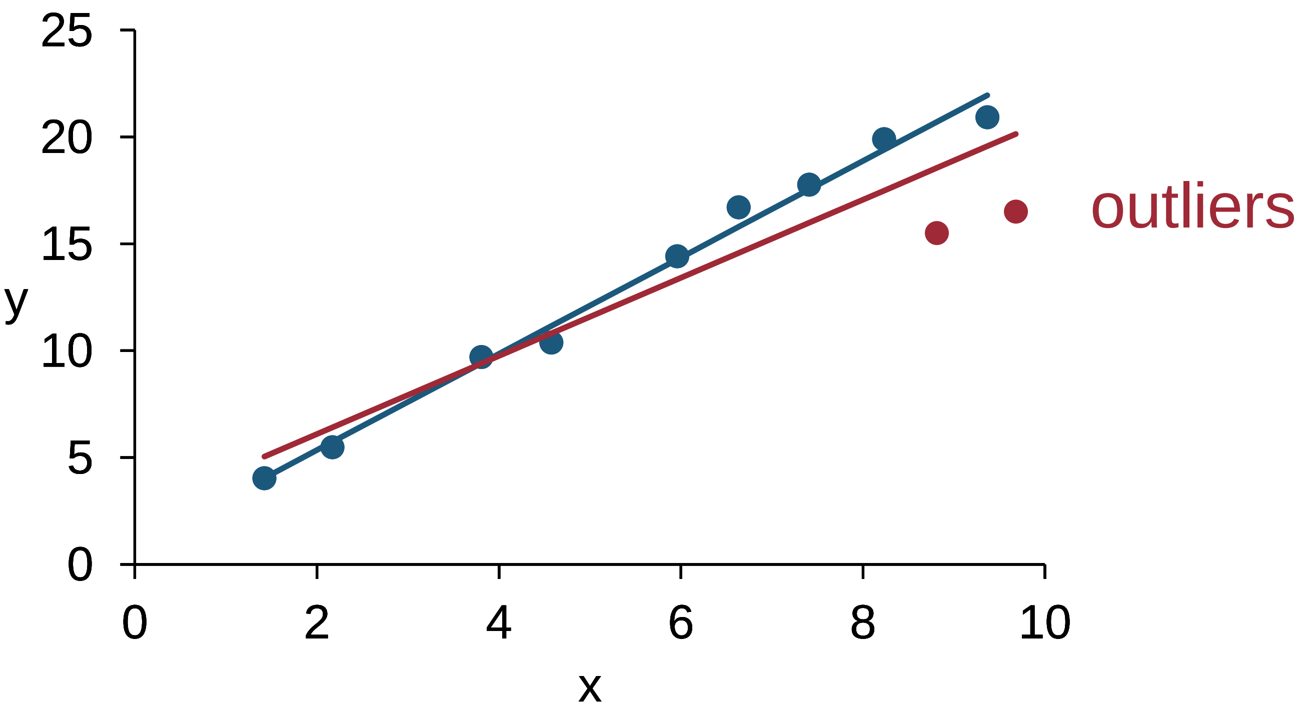
Leow Wee Kheng
CS4243 Computer Vision and Pattern Recognition

Robust Methods

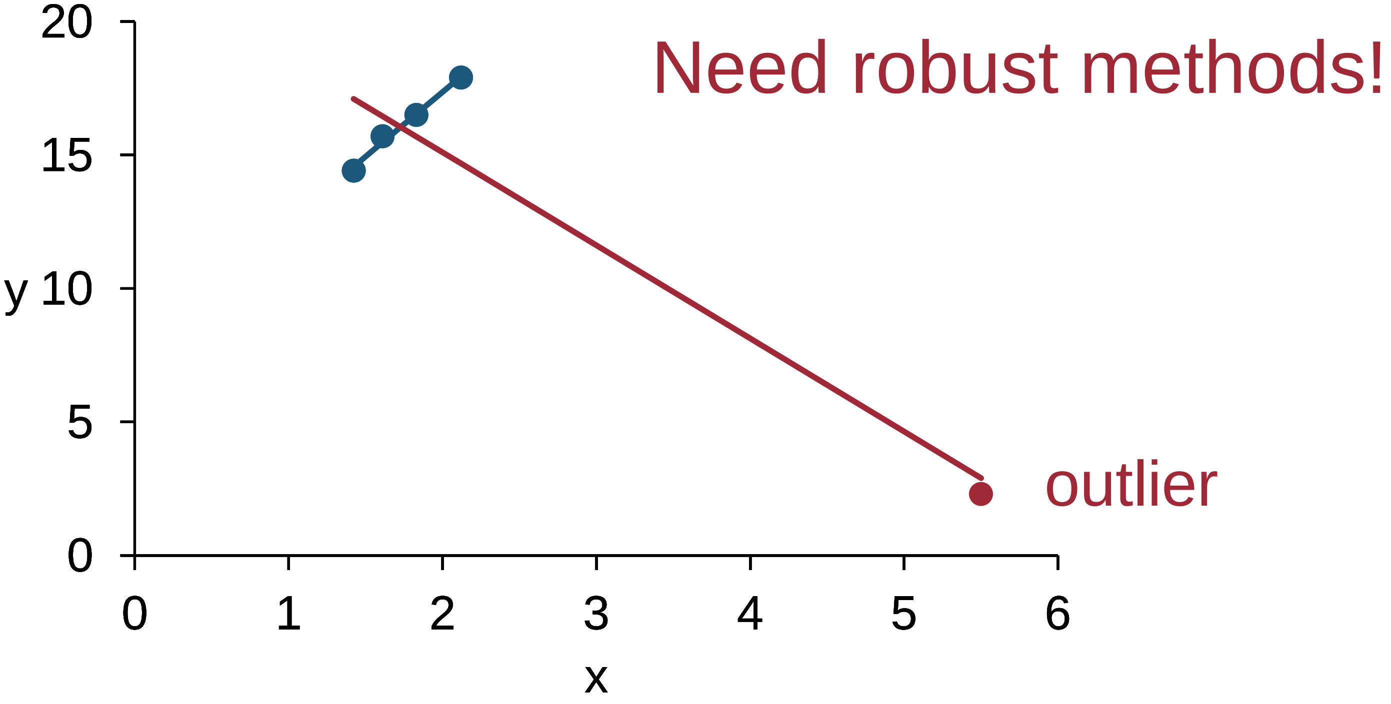
Consider this data set…

- Fitting a line to blue points give blue line.
- Outliers cause fitting error (red line).



Consider this data set…

- Fitting a line to blue points give blue line.
- Single outlier can cause serious error!



Robust Methods

Two basic approaches:

- Make error functions immune to outliers.
- Detect and discard outliers.

Exercise

- Consider this set of number:

1, 2, 3, 4, 5, 6, 7

- What is the mean?
- What is the median?
- Now, insert an outlier

1, 2, 3, 4, 5, 6, 7,14

- What is the mean, error of mean?
- What is the median, error of median?
- Which one robust, mean or median?

Robust Statistics

- Mean square error is not robust

$$E = \frac{1}{N}\sum_i \left(f(\mathbf{x}_i) - y_i\right)^2$$

- Median square error is robust

$$E = \operatorname*{median}_i \left(f(\mathbf{x}_i) - y_i\right)^2$$

  - Can potentially tolerate up to *N* / 2 outliers!

Robust Error Functions

- Not adversely affected by outliers.
- Consider least square fitting

$$E = \sum_i (f(x_i; \mathbf{a}) - y_i)^2$$

  - Want to find parameters **a** that minimise *E*.
  - Error $r_{i,\mathbf{a}} = f(x_i; \mathbf{a}) - y_i$ is also called residue.
  - Error function can be rewritten as

$$E = \sum_i r_{i,\mathbf{a}}^2$$

  - Square of r increases very fast with r. Not robust.

- Robust functions $\rho(r)$ increase slowly with r.

- Properties of $\rho(r)$:
  - Global minimum: $\rho(r) = 0$ when $r = 0$
  - Positive: $\rho(r) \geq 0$
  - Symmetric: $\rho(r) = \rho(-r)$
  - Monotonically increasing: $\rho(r) \geq \rho(s)$ iff $r \geq s$

Robust Error Functions

- Beaton and Tukey

$$\rho(r) = \begin{cases} \frac{1}{6}a^2\left[1 - \left(1 - \left(\frac{r}{a}\right)^2\right)^3\right] & \text{if } |r| \le a \\ \boxed{\frac{1}{6}a^2 \qquad \text{otherwise}} \end{cases}$$

fixed for large r

- Cauchy

$$\rho(r) = \frac{a^2}{2}\boxed{\log\left[1 + \left(\frac{r}{a}\right)^2\right]}$$

increases slowly with r

Robust Error Functions

- Huber

$$\rho(r) = \begin{cases} \frac{1}{2}r^2 & \text{if } |r| \le a \\ \boxed{\frac{1}{2}a(2|r| - a) \quad \text{otherwise}} \end{cases}$$

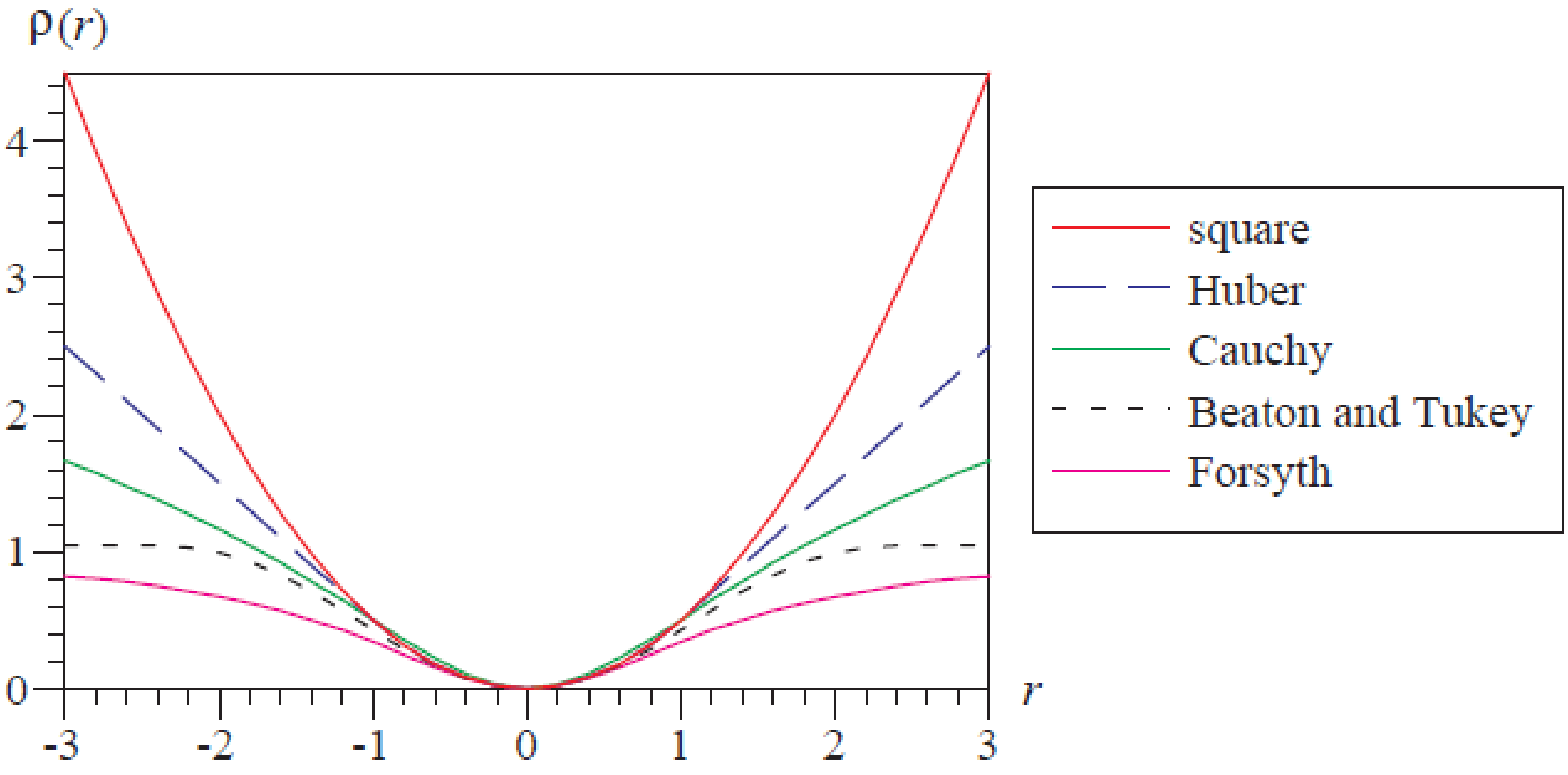
increases
linearly with r

- Forsyth

$$\rho(r) = \boxed{\frac{r^2}{a^2 + r^2}}$$

almost fixed
for large r

Robust Error Functions



Random Sample Consensus

- RANSAC: Random Sample Consensus
- Robust approach for solving many problems.
- Try to identify outliers.
- Basic ideas
  - RANSAC wraps over target algorithm.
  - Start with small random subset of data points.
  - Run target algorithm on subset.
  - Iteratively add consistent data points to subset.

- Target algorithm has two things
  - Model *M*: function parameters **a**, etc.
  - Error measure *E*: need not be robust

- Example: least square fit

error

model

$$E = \sum_i \left(f(x_i; \mathbf{a}) - y_i\right)^2$$

- Example: affine transformation

error

model

$$E = \sum_{i=1}^{n} \|\mathbf{p}_i' - \mathbf{A}\,\mathbf{p}_i\|^2$$

RANSAC

- Given N data points.
- Initialise best model $M^* \leftarrow \varnothing$, error $E(M^*) \leftarrow \infty$.
- Repeat for k iterations:
  1. Randomly select subset S of $m < N$ data points.
  2. Run target algorithm on S to determine M.
  3. For each data point $\mathbf{x}$ not in S
     - If error of M on $\mathbf{x} < $ tolerance τ, add $\mathbf{x}$ to S.
  4. If $|S| >$ threshold Γ
     - Use S to determine new M.
     - If $E(M) < E(M^*)$, update $M^* \leftarrow M$, $S^* \leftarrow S$.

consistent set

- RANSAC has 3 parameters
  - Error tolerance τ
    - Dependent on the expected error of fitting inliers.
  - Size threshold Γ
    - Dependent on the model.
    - Should be large enough to have enough inliers.
  - Number of iterations k
    - Should be large enough to get good model.
    - Can result in large execution time.

Minimum Subset Random Sampling

- Use robust error function and random sampling.
- Basic algorithm

1. Initialise best model $M^* \leftarrow \varnothing$, error $E(M^*) \leftarrow \infty$.
2. Randomly select k subsets of m data points.
3. For each subset S of m data points
   - Determine model M that best fits data points.
   - Compute error E of applying M on all N data points.
   - If $E(M) < E(M^*)$, update $M^* \leftarrow M$, $S^* \leftarrow S$.

must be immune to outliers

- Algo complexity: $O(kN)$

How to be immune to outliers?

- Use robust statistics
  - Median ($r_1, \ldots, r_N$)
- Use error function with fixed value for larger *r*
  - Beaton and Tukey

$$\rho(r) = \begin{cases} \frac{1}{6}a^2\left[1 - \left(1 - \left(\frac{r}{a}\right)^2\right)^3\right] & \text{if } |r| \le a \\ \boxed{\frac{1}{6}a^2 \qquad \text{otherwise}} \end{cases}$$

fixed for large r

- Use truncated error function

$$\rho(r) = \begin{cases} e(r) & \text{if } r < a \\ e(a) & \text{otherwise} \end{cases}$$

  - $e(r)$ is any error function
  - $\rho(r)$ is fixed for large r

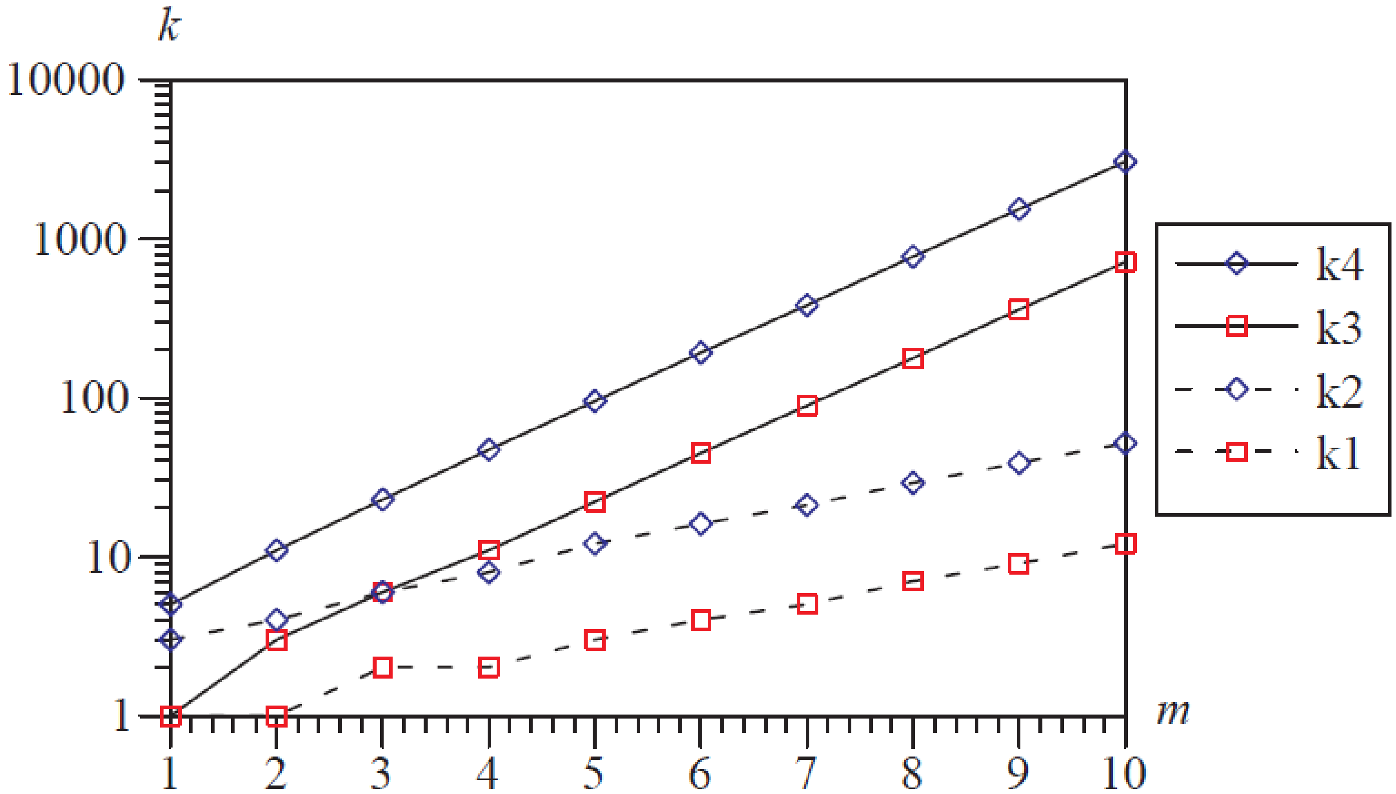
Selection of *k*

- Max value of *k* is very large: ${}^{N}C_{M}$
- In practice, smaller value of *k* is possible.
- Analysis
  - ε = prob. randomly selected data point is inlier
  - ε = fraction of inliers in data set
  - ε^m = prob. all *m* data points in subset are inliers
  - *s* = prob. at least 1 subset has only inliers

$$s = 1 - (1 - \varepsilon^m)^k$$

  - Then,

$$k = \frac{\ln(1 - s)}{\ln(1 - \varepsilon^m)}$$

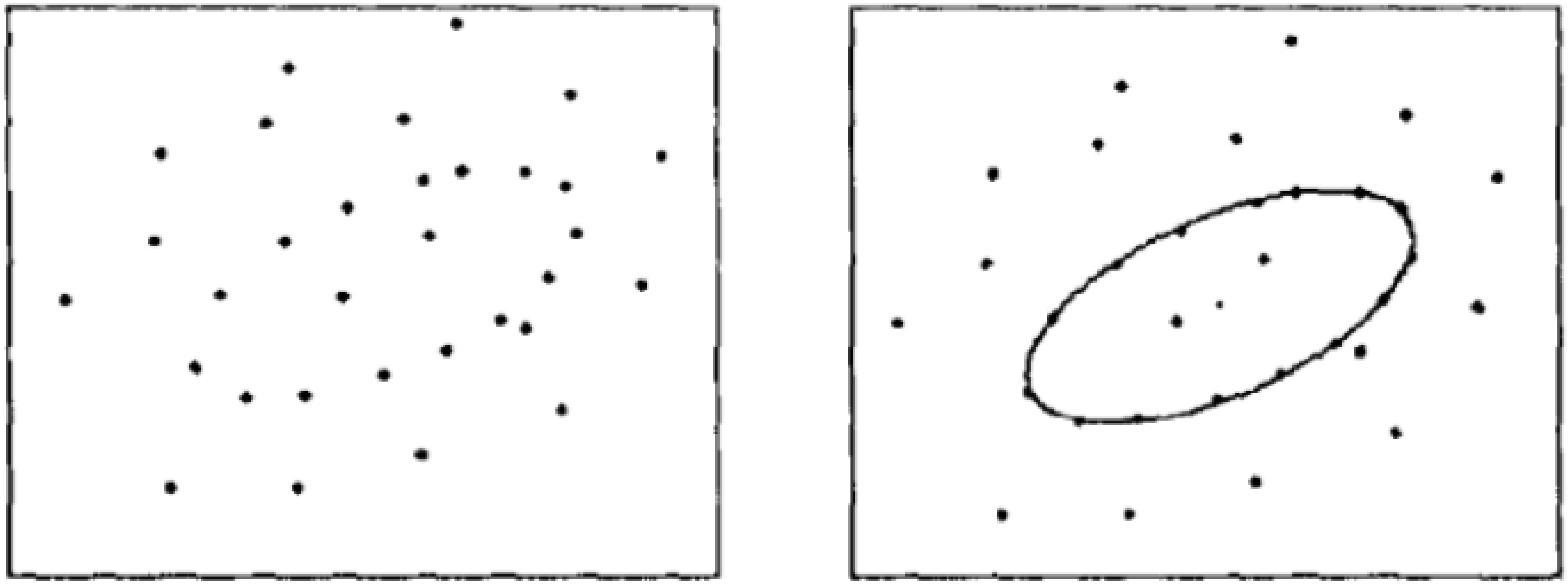


	s = 0.5	*s* = 0.95
ε = 0.75	k1	k2
ε = 0.5	k3	k4

small *k*

Example

- Fit ellipse to data points [RL93]



Summary

- Outliers can adversely affect algorithm's error.
- Robust error functions are immune to outliers.
- RANSAC
  - Robust method for finding consistent set of inliers.
- Minimum subset random sampling
  - Use robust error function and random sampling to speed out random sampling.

Further Reading

- Robust parameter estimation
  - [Ste99], [MMRK91]
- RANSAC
  - [FB81], [FP03] Section 15.5, 15.6.
- Minimal Subset Random Sampling
  - [RL93]

References

- M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with apphcatlons to image analysis and automated cartography. *Communications of ACM, 24(6):381–395, 1981.*
- D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach.* Pearson Education, 2003.
- P. Meer, D. Mintz, A. Rosenfeld, and D. Kim. Robust regression methods in computer vision: A review. *Int. Journal of Computer Vision, 6(1):59–70, 1991.*
- G. Roth and M. D. Levine. Extracting geometric primitives. *CVGIP: Image Understanding, 58(1):1–22, 1993.*
- C. V. Stewart. Robust parameter estimation in computer vision. *SIAM Review, 41(3):513–537, 1999.*