**Example Problem 1**

How to fit a line to a set of points \( \{(x_i, y_i)\} \)?

\[
\text{Method 1:}
\]

Model is \( y_i = mx_i + b + \epsilon_i \)

\[
\Rightarrow \text{minimize } \sum_i (y_i - mx_i - b)^2
\]

\[
\begin{align*}
\text{To minimize } & \sum_i (y_i - mx_i - b)^2 \\
\text{Take } & \frac{\partial}{\partial m}, \frac{\partial}{\partial b} \text{ and set to } 0. \\
\Rightarrow & \sum_i x_i (y_i - mx_i - b) = 0 \\
\quad & \sum_i (y_i - mx_i - b) = 0
\end{align*}
\]

and solve for \( m, b \).

**Method 1**

"Total Least Square"

Errors in \( y \) vs. Errors perpendicular to line

- \( r \) is perpendicular distance from point \( (x_i, y_i) \)
- \( |x_i \cos \theta + y_i \sin \theta - r| \) is distance from \( (x_i, y_i) \) to line

**Method 2**

"Least squares"

Errors in \( x \) only

**Sum of squared errors perpendicular to model**

What is \( \theta, r \) that minimizes it?
Example Problem 2: "vanishing point detection"

Suppose you have a set $S$ of edges $\{ (x_i, y_i; \Theta_i) \}$,
where $\cos \Theta_i$ is 1 to $
$
Estimate the intersection point $(x_v, y_v)$
of lines defined by these edges:

If $(x_v, y_v)$ is on the line $(x, y; \Theta)$,
then $(x_v - x, y_v - y) \cdot (\cos \Theta, \sin \Theta) = 0$

Minimize

$$\sum \frac{1}{2} (x_i - x_v, y_i - y_v) \cdot (\cos \Theta_i, \sin \Theta_i)^2$$

Minimize

$$\sum \frac{1}{2} (x_i - x_v, y_i - y_v) \cdot (e_1, e_2)^2$$

when $(x_i, y_i; \Theta_i)$ are given.

This is of the form:

$$\text{minimize} \frac{1}{2} \sum \| A \tilde{x} - \tilde{b} \|^2$$

where $A, \tilde{b}$ are given.

Later we will look at solution to
this general problem.

Limitations of Least Squares Methods

- inliers - model + noise
- outliers - not model
don't want to use these samples

Reducing sensitivity to outliers

How to generalize each point based on its

- distance squared
- distance

- distance > $\tau$ \{ less sensitive to outliers

Hough Transform

for each $(x_i, y_i)$
for each $\Theta$
\begin{align*}
    r &= \text{round} \left( x_i \cos \Theta + y_i \sin \Theta \right) \\
\end{align*}

\} \text{ vote for } (\Theta, r)

\begin{align*}
    \text{count number of samples that have distance } > \tau \text{ from line} \\
    \text{return } (r, \Theta) \text{ with smallest count} \\
    \text{count number of samples that have distance } \leq \tau \text{ from line} \\
    \text{return } (r, \Theta) \text{ with largest count}
\end{align*}

This is basically the
second formulation
on the previous
slide!
Hough Transform

for each \((x_i; y_i)\)
    for each \(\theta\)
        \(r = \text{round} \left( \frac{y_i}{\cos \theta} + \frac{x_i}{\sin \theta} \right)\)
        vote for \((\theta, r)\)

\(r\)
\(\theta\)
\((\theta + \pi, -r)\)

Hough can detect multiple models/peaks

\(\checkmark\) Quantization of \((r, \theta)\) is a problem, especially if there is noise.
Peaks might be difficult to detect.

Hough transform for vanishing points

Suppose you have a set \(S\) of edges \((x_i; y_i; \theta_i)\).

Find \((x_v, y_v)\).

Hough transform for vanishing points (1)

Suppose you have a set \(S\) of edges \((x_i; y_i; \theta_i)\).

for each \((x_i; y_i; \theta_i)\)
    for each \(t\)
        vote for \((x_i, y_i) + t(-\sin \theta_i, \cos \theta_i)\)
    pick \((x_v, y_v)\) with the most votes

Hough transform for vanishing points (2)

Vanishing points are often not within the field of view.

Vote on the unit sphere instead.

Need to tile the sphere.

(Details omitted.)

Hough Transform (general)

Two flavours
- each feature votes for multiple models
- multiple features are combined and vote one model

Either way, choose model that gets the most votes.

Hough often works well for 2-d model spaces e.g (\((r, \theta)\), \((x_v, y_v)\))

But for higher dimensional models, it doesn't work well.
We will see 8-d models (and higher!) soon.
Least Squares
- use all data to fit model
  😞 sensitive to outliers
  😊 good fit if most points are inliers

RANSAC
- use minimal data to fit model
  😊 increase chances of no outliers
  😞 sensitive to noise

RANSAC (Random Sample Consensus)
- repeat
  • randomly sample n points = 1 "trial"
    (minimum needed to fit the model e.g. n=2)
  • fit a model for this trial
  • examine all other points (N-n) and see
    how many are within distance T1 from the
    model (called the "consensus set")
  • increment counter
    until (consensus set > T2) or (counter = numTrials)
- re-fit model using largest consensus set
  and return

RANSAC - Examples

\[ y = mx + b \]

Finding a vanishing point
(pairs of edges define a plane)
(pairs of points define a line)

RANSAC
Let \( p \) be probability that
a sample is an inlier.
The probability that at least one of
the \( n \) samples contains an outlier is
\[ 1 - p^n \]

• What is \( (1-p^n)^{numTrials} \) ?

• How might you estimate \( p \) ?