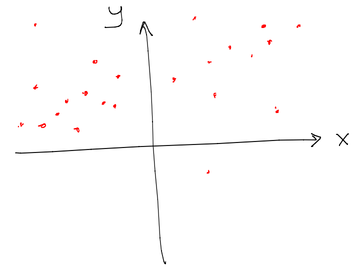
lecture 15

model fitting methods

- least squares
- Hough transform
- RANSAC

Example Problem 1:

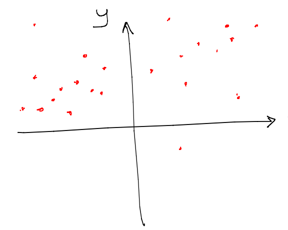
How to fit a line to a set of points $\{(x_i\, y_i)\}$?



Method 1:

Model is $y_i = m x_i + b + n_i$

$$\Rightarrow \underset{\text{over } m,b}{\text{minimize}} \sum_i (y_i - m x_i - b)^2$$



To minimize $\sum_i (y_i - m x_i - b)^2$

Take $\frac{\partial}{\partial m}$, $\frac{\partial}{\partial b}$ and set to 0.

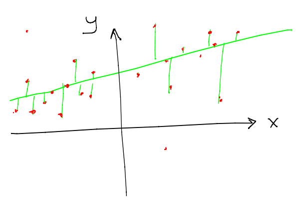
$$\sum_i (y_i - m x_i - b)\, x_i = 0$$

$$\sum_i (y_i - m x_i - b) = 0$$

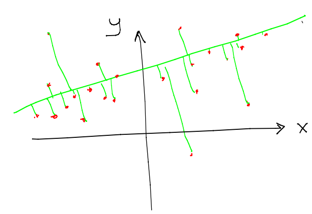
$$\Rightarrow \sum_i \underbrace{\begin{bmatrix} x_i^2 & x_i \\ x_i & 1 \end{bmatrix}}_{2\times 2} \begin{bmatrix} m \\ b \end{bmatrix} = \sum_i \underbrace{\begin{bmatrix} x_i y_i \\ y_i \end{bmatrix}}_{2\times 1}$$

and solve for m, b.

Method 1		Method 2 "Total Least Squares"
Errors in y direction only	vs	Errors perpendicular to line

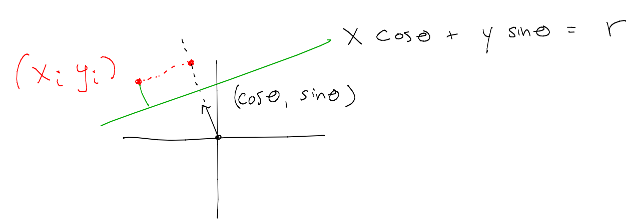


$y = mx + b$



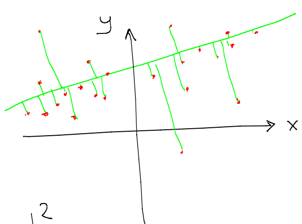
$x\cos\theta + y\sin\theta = r$

How to measure perpendicular distance?



- r is perpendicular distance from origin to the line.
- $|x_i \cos\theta + y_i \sin\theta - r|$ is distance from $(x_i\, y_i)$ to line

Sum of squared errors perpendicular to model



$$\sum_i |x_i \cos\theta + y_i \sin\theta - r|^2$$

What is θ, r that minimizes it?

$$\frac{\partial}{\partial r} \sum_{i=1}^{n} |x_i \cos\theta + y_i \sin\theta - r|^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} (x_i \cos\theta + y_i \sin\theta - r) = 0$$

$$\Rightarrow \cos\theta \frac{\sum_i x_i}{n} + \sin\theta \frac{\sum_i y_i}{n} = r$$

$$\Rightarrow (\bar{x}, \bar{y}) = \left(\frac{\sum x_i}{n}, \frac{\sum y_i}{n} \right)$$

is on the best fit line.

$$\cos\theta\, \bar{x} + \sin\theta\, \bar{y} = r$$

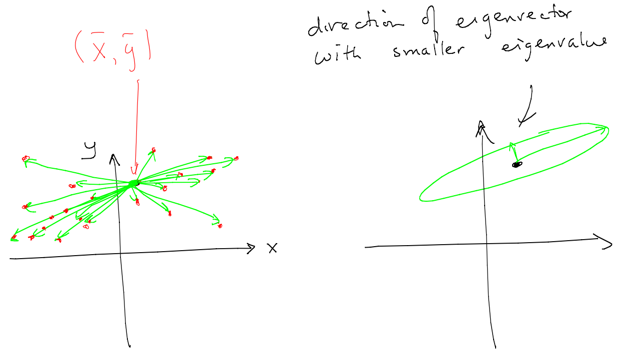
Minimize $\sum_i (\cos\theta\, x_i + \sin\theta\, y_i - r)^2$

$$\equiv \text{minimize} \sum_i \left(\cos\theta\, (x_i - \bar{x}) + \sin\theta\, (y_i - \bar{y})\right)^2$$

$$\begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} \underbrace{\begin{bmatrix} \leftarrow & x_i - \bar{x} & \rightarrow \\ \leftarrow & y_i - \bar{y} & \rightarrow \end{bmatrix}}_{2\times n} \underbrace{\begin{bmatrix} \uparrow & \uparrow \\ x_i - \bar{x} & y_i - \bar{y} \\ \downarrow & \downarrow \end{bmatrix}}_{n\times 2} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$\sum_i \begin{bmatrix} (x_i - \bar{x})^2 & (x_i - \bar{x})(y_i - \bar{y}) \\ (x_i - \bar{x})(y_i - \bar{y}) & (y_i - \bar{y})^2 \end{bmatrix}$$

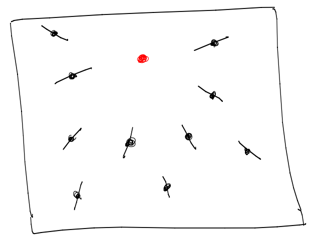
Find eigenvector of 2x2 matrix with smaller eigenvalue.



Example Problem 2 : "vanishing point detection"

Suppose you have a set of edges $(x_i\ y_i\ \theta_i)$

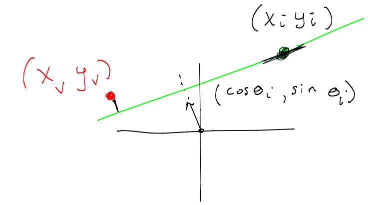
where $\cos\theta_i\ \sin\theta_i$ is $\perp$ to edge.



Estimate the intersection point (x_v, y_v) of lines defined by these edges.

If (x_v, y_v) is on the line (x_i, y_i, θ) then

$$(x_i - x_v,\ y_i - y_v) \cdot (\cos\theta_i, \sin\theta_i) = 0$$



$\therefore$ Minimize

$$\sum_i \left((x_i - x_v,\ y_i - y_v) \cdot (\cos\theta_i, \sin\theta_i) \right)^2$$

$\therefore$ Minimize

$$\sum_i \left((x_i - x_v,\ y_i - y_v) \cdot (\cos\theta_i, \sin\theta_i) \right)^2$$

where $(x_i\ y_i\ \theta_i)$ are given.

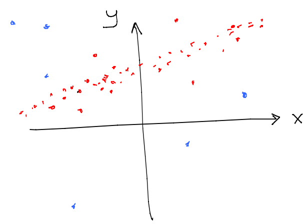
This is of the form:

$$\text{minimize} \quad \| \underbrace{\vec{A}}_{2\times 2} \underbrace{\vec{x}}_{2\times 1} - \underbrace{\vec{b}}_{2\times 1} \|^2$$

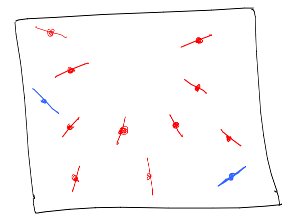
where A, b are given.

Later we will look at solution to this general problem.

Limitations of Least Squares Methods



fitting a line

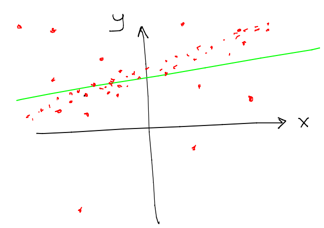


finding a vanishing point

inliers – model + noise

$\left.\begin{array}{l}\text{outliers – not model}\end{array}\right\}$ don't want to use these samples

Reducing sensitivity to Outliers

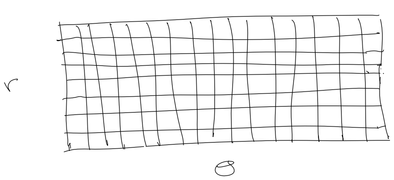


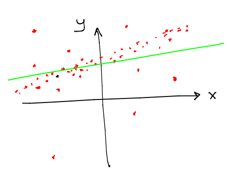
How to penalize each point based on its distance from a candidate line?

- distance squared
- distance
- distance $> \tau$

} less sensitive to outliers

Voting Method





for each (r, θ)
  Count number of samples that have distance $> \tau$ from line
return (r, θ) with smallest count

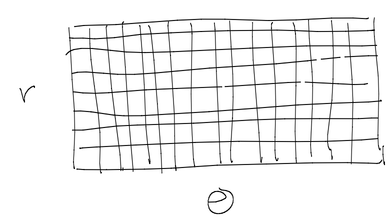
Count number of samples that have distance $> \tau$ from line
return (r, θ) with smallest count

equivalent

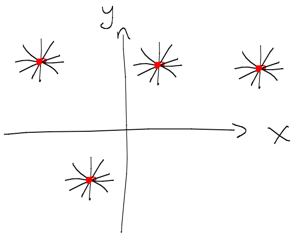
Count number of samples that have distance $\leq \tau$ from line
return (r, θ) with largest count

Hough Transform

for each $(x_i\ y_i)$
  for each θ {
    $r = \text{round}(x_i \cos\theta + y_i \sin\theta)$
    vote for (θ, r)
  }
}



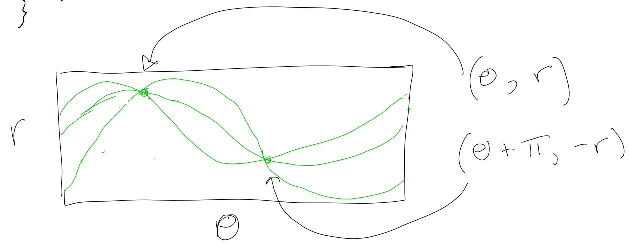
This is basically the second formulation on the previous slide!

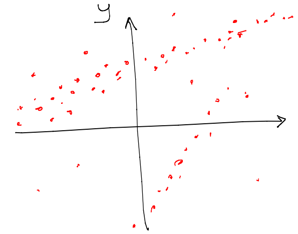


The line through the top three points gets 3 votes.
(There are also 3 lines that get 2 votes.)

Hough Transform

```
for each (x_i y_i)
  for each θ {
    r = round (x_i cos θ + y_i sin θ)
    vote for (θ, r)
  }
}
```

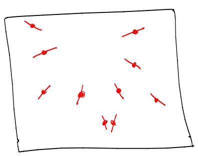




✓ Hough can detect multiple models/peaks

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

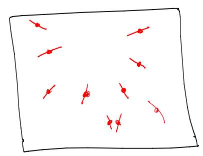
Hough transform for vanishing points



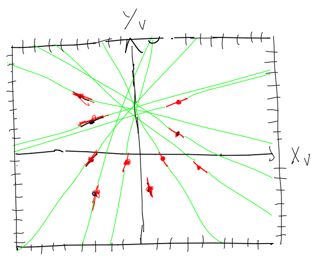
Suppose you have a set of edges $(x_i\ y_i, \theta_i)$.

Find $(x_v\ y_v)$.

Hough transform for vanishing points (1)

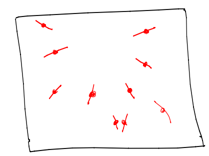


Suppose you have a set of edges $(x_i\ y_i, \theta_i)$.

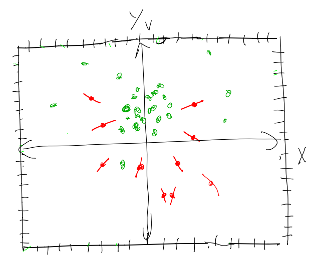



```
for each (x_i, y_i, θ_i)
  for each t {
    vote for
    (x_i y_i) + t · (−sin θ_i, cos θ_i)
  }
pick (x_v y_v) with the most votes
```

Hough transform for vanishing points (2)



Suppose you have a set of edges $(x_i\ y_i, \theta_i)$.

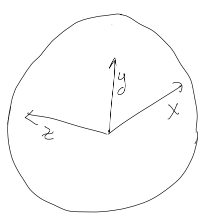



```
for each pair of edges
  (x_i, y_i, θ_i) (x_j, y_j, θ_j)
  Compute intersection
  (x_v y_v) of lines.
  and vote for it
}
pick (x_v y_v) with the most votes
```

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 eg (r, θ), (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

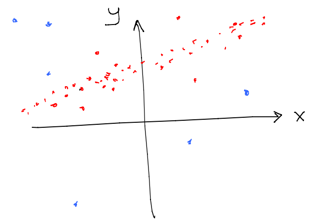
- use minimal data to fit model
  - ☺ increase chances of no outliers
  - ☹ sensitive to noise

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 τ_1 from the model (called the "consensus set")
  - increment counter

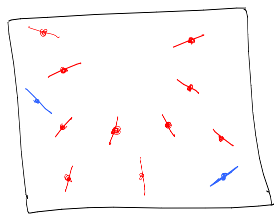
  until (consensus set $> \tau_2$) or (counter == numTrials)
- refit model using largest consensus set and return

RANSAC - Examples



fitting a line

(pairs of points define a line)



finding a vanishing point

(pairs of edges define a vanishing point)

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 ?