Graduate Computer Vision

CS670

Unit 2: Probability, Statistics, Supervised Learning, and Intro to Classification

Erik Learned-Miller

Today

- Some central problems of computer vision
- Review of basic probability:
  - Conditional probability, joint probability, marginal probabilities, likelihoods, priors, posteriors, and Bayes' rule.
- Learning from example: Supervised Learning
- Simple features
- More Matlab

Central Tasks of Computer Vision

- ?

Central Tasks of Computer Vision

- Navigation
  - Obstacle avoidance
  - Memory (must be efficient!)
  - Recognition of previously seen locations.
- Detection of motion (why?)
- Estimation of depth
- Individual object recognition (object I've seen before)
- Category recognition: Is it an apple?

Subtleties of Recognition

Wikimedia Commons: Photo by Muhammad Mahdi Karim

Subtleties of Recognition

Subtleties of Recognition

Subtleties of Recognition

Wikimedia commons: public domain

Recognition

- Before you can recognize something
  - You have to figure out whether there is anything there to recognize at all
  - You have to find the thing you want to recognize
- Dilemma: How to search for something to recognize if you don't know what you're looking for?

Review of Basic Probability

- Sample space, event space, random experiments
  - Covered in on-line document. I will not review.
- **Calculating probability of events**.
  - Consider an unfair die with probabilities: {1/8, 1/8, 1/8, 1/8, 1/4, 1/4}
  - Compute probability that roll is greater than 3 or odd.

Computing Probabilities

1. Enumerate all primitive events which satisfy criterion: {1, 3, 4, 5, 6}.
2. Add probabilities of primitive events: $1/8 + 1/8 + 1/8 + ¼ + ¼ = 7/8$.

Joint Probability

- Let A and B be events.
- The probability that *both* event A and event B occurred *on the same trial of an experiment.*

$$P(A, B)$$

Joint Probability

- Let X and Y be random variables, where X can take on the values {0, 1} and Y can take on the values {1,2,3,4,5,6}.
- The joint probability that X=1 and Y=3 is written

$$P(X = 1, Y = 3)$$

Notation and Types

- Let X and Y be random variables, where X can take on the values {0, 1} and Y can take on the values {1,2,3,4,5,6}.
- What is the *type* of each of these?

$$P(X)$$

$$P(X, Y = 5)$$

$$P(X, Y)$$

$$P(X = 1, Y)$$

Marginal Probability

Let X and Y be two random variables, such as the outcomes of a blue die and a red die which are tossed together. If we are given the probabilities of all events $P(X = x, Y = y)$ in the joint sample space, then we can compute the probability of events involving only a single random variable, such as $P(Y = 3)$, through a process known as *marginalization*. In particular, we can say that

$$\begin{aligned}
& P(Y = 3) \\
= \quad & P(X = 1, Y = 3) + P(X = 2, Y = 3) + P(X = 3, Y = 3) \\
& \quad + P(X = 4, Y = 3) + P(X = 5, Y = 3) + P(X = 6, Y = 3) \\
= \quad & \sum_{x=1}^{6} P(X = x, Y = 3).
\end{aligned}$$

Conditional Probability

When A and B are events on an event space S, we read $P(A|B)$ as *the probability of event A given that the event B has occurred on the same trial*, or more succinctly, *the probability of A given B*. This is also referred to as the *conditional probability* of A given B.

$$P(A, B) = P(A|B)P(B).$$

Computing Conditional Probabilities: Approach 1

$$P(A, B) = P(A|B)P(B).$$

yields

$$P(A|B) = \frac{P(A, B)}{P(B)}.$$

Computing Conditionals

Example 6. Assume an unfair die with probabilities as in the example above: $P(1) = \frac{1}{16}, P(2) = \frac{1}{16}, P(3) = \frac{1}{16}, P(4) = \frac{1}{16}, P(5) = \frac{1}{4}, P(6) = \frac{1}{2}$. Let A be the event that a die roll is less than 5. Let B be the event that the die roll is prime. Compute $P(A|B)$.

Bayes' Rule

- Central to artificial intelligence
- Allows us to "reverse the direction of conditioning"
  - We can compute $P(A|B)$ from $P(B|A)$, $P(A)$, and $P(B)$.
  - Why is this useful?

Bayes' Rule

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A).$$

Bayes' Rule

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A).$$

By simply dividing both sides of the latter equation by $P(B)$, we have

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)},$$

or, alternatively, dividing by $P(A)$, we have

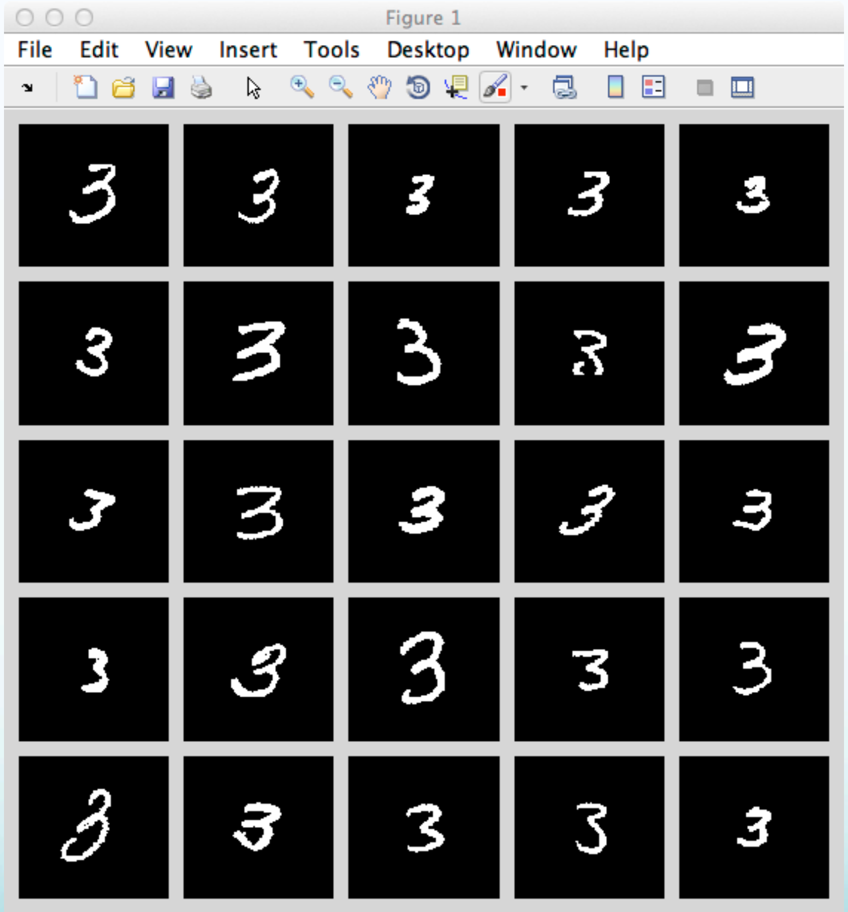
$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

Example 7 from reading...

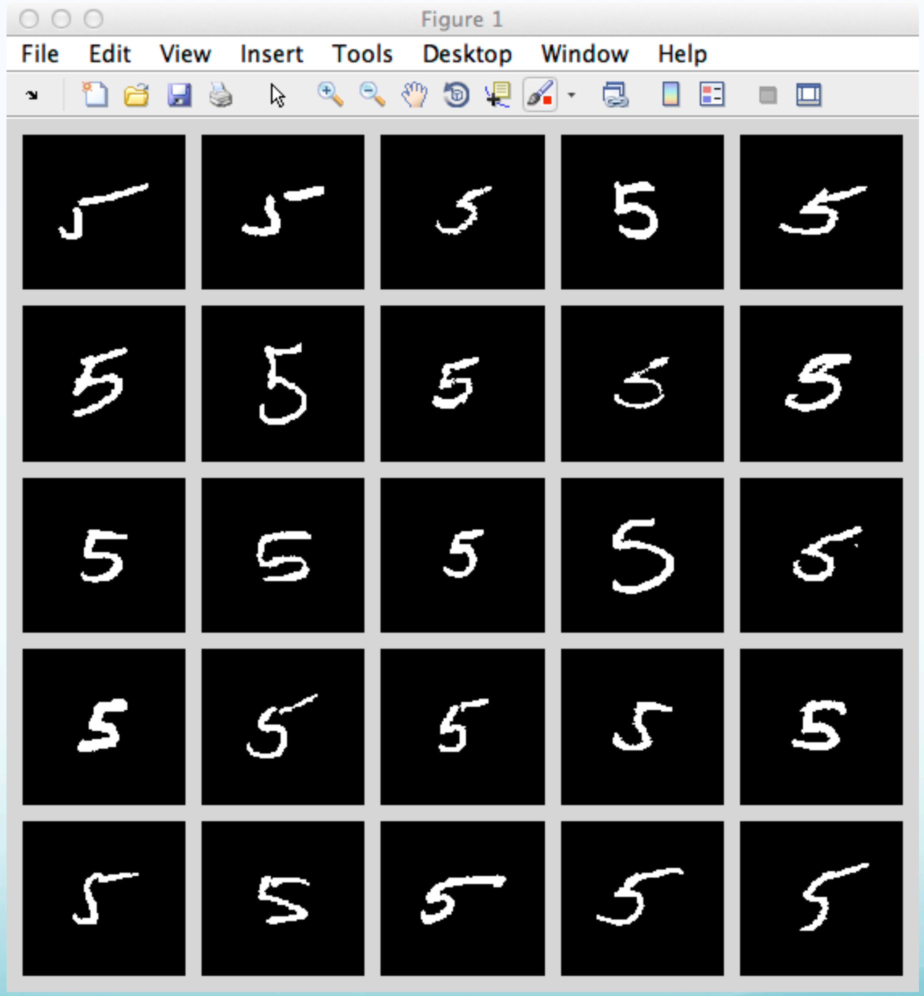
First approach to recognition

- Easy problem: handwritten digit recognition
  - Distinguish between 3's and 5's.
- Simple images: binary
- Digits have been written on a black background
- Digits are approximately centered and scaled to be close to a uniform size

NIST Handwritten Digits



NIST Handwritten Digits



Strategies for Recognition

- ?

Strategies for Recognition

- Traditional approach (1960’s, 1970’s, ...)
  - Define some “obvious” features of 3’s and 5’s.
  - Look for them in the images.
  - Problems:
    - “obvious” features often are absent
    - Not clear how to write algorithms to find these “obvious” features.

Strategy 2: Learn from Examples

- Given: Some examples of 3's, some examples of 5's.
- Find: a rule which tends to classify images of 3's as 3's and images of 5's as 5's on training data.
- Supervised learning: *formalization of learning from example*

Supervised learning

- Training Data (with labels):

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)$$

- Test Data:

$$(\mathbf{x}_{n+1}, \mathbf{x}_{n+2}, \ldots, \mathbf{x}_{n+m})$$

Supervised learning

- Test data drawn from some joint distribution

$$p(\mathbf{X}, Y)$$

- What if *training data* are drawn from a completely different distribution?

Supervised learning

- Today's approach:
  - Split training samples into two groups, those for which Y=1 and for which Y=2.
  - Estimate p(**X**|Y=1), and p(**X**|Y=2).
  - Also estimate p(Y).
  - Now use Bayes' rule to calculate p(Y|**X**).

How to estimate p(**X**|Y)

- Impracticality of doing this for full images.
- Must rely on simple features.

Matlab