Carnegie Mellon
THE ROBOTICS INSTITUTE

Expectations and Covariances

Robot Localization and Mapping
16-833

Michael Kaess

February 3+8, 2021

Slides courtesy of Ryan Eustice

Probabilistic State Estimation

- Uncertain observations
  - Sensor noise & non-idealities
- Uncertain beliefs
  - Derived from sensor observations
  - Approximate algorithms
- *Probabilistic State Estimation*
  - Identify the quantities (state variables) we care about.
  - Determine probability for every possible simultaneous assignment

Representing State

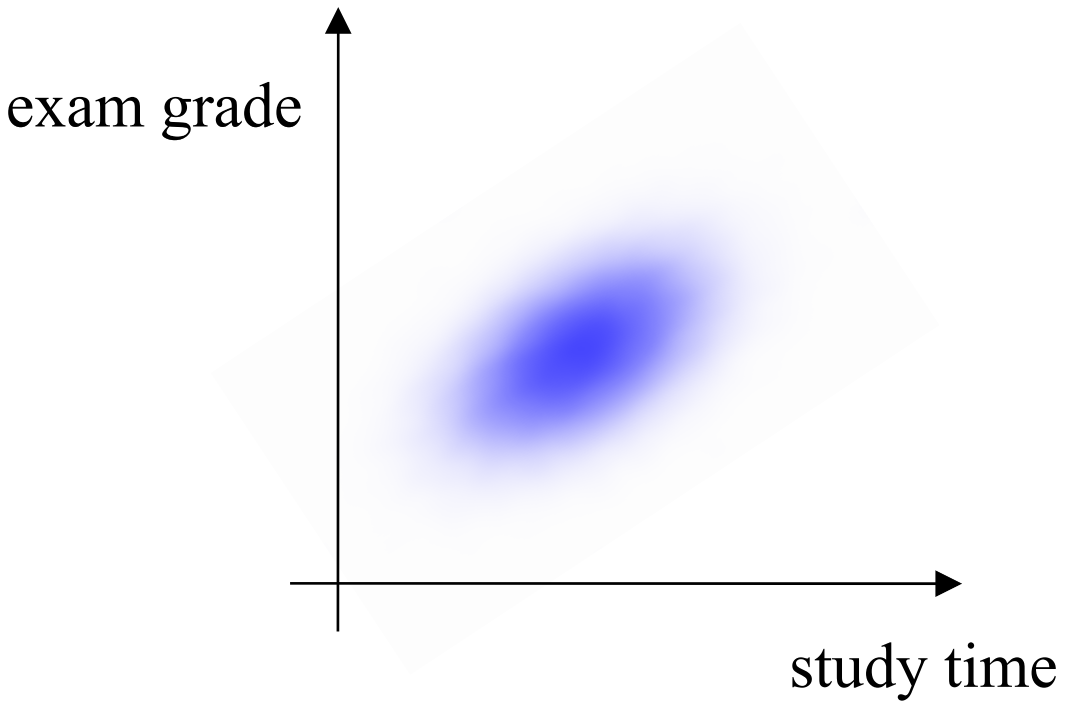
- Represent everything we need to know in terms of a vector of quantities
  - "State vector"
  - Usually continuous-valued in this course

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}$$

- The "meaning" of the variables is up to us
  - e.g., index 7 is the temperature in Seattle.
  - Bookkeeping work for us.

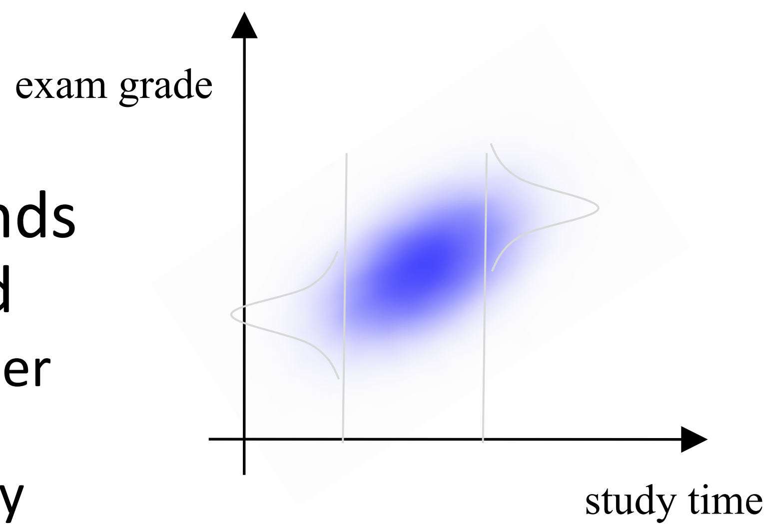
Representing Uncertainty

- In principle, distribution of unknown quantities can be arbitrary



Correlations

- Estimates of variables tend to become correlated over time
  - Observation: Study time is 4 hours
  - Belief about study time *and* exam grade are affected
- Distribution of exam grade depends on study time: two are correlated
  - We'll look at correlations closer later today
  - The data does not necessarily imply any *causal* relationship.



Probability Basics

Discrete Probability	Continuous Probability
P(x) = Probability of event occurring	p(x) = Probability *density* at x
$Prob(x) = P(x)$	$Prob(x) = 0$
$0 \leq P(x) \leq 1$	$0 \leq p(x) < \infty$
$\sum_{-\infty}^{\infty} P(x) = 1$	$\int_{-\infty}^{\infty} p(x)dx = 1$

Probability Basics: Expectation

- Weighted average according to probability

$$\mu_x = E[x] = \int_{-\infty}^{\infty} x p(x) dx$$

- Basic properties of expectation

$$E[\alpha] = \alpha$$

$$E[\alpha x] = \alpha E[x]$$

$$E[\alpha + x] = \alpha + E[x]$$

$$E[x + y] = E[x] + E[y]$$

Joint Expectation

$$E[xy] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, p(x,y)dxdy$$

- Uncorrelated: $E[xy] = E[x]E[y]$
  - Independence → Uncorrelated
  - Uncorrelated $\not\rightarrow$ Independence
    - e.g.

$$p(x,y) = \tfrac{1}{4}\delta(x,y-1) + \tfrac{1}{4}\delta(x,y+1) + \tfrac{1}{4}\delta(x-1,y) + \tfrac{1}{4}\delta(x+1,y)$$

- Conditional Expectation: $E[x|y] = \int_{-\infty}^{\infty} x\, p(x|y)dx$
  - $E[x|y] = E[x]$ implies neither independence nor uncorrelatedness
    - e.g. $p(x,y) = \dfrac{1}{3}\delta(x,y+1) + \dfrac{1}{3}\delta(x+1,y) + \dfrac{1}{3}\delta(x-1,y)$

Variance & Covariance

- Average squared deviation from the mean.

- (Auto) covariance
  - Scalar: $$\sigma_x^2 = E[(x - E[x])^2]$$
  - Vector: $$\Sigma_{\mathbf{xx}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^\top]$$

- (Cross) covariance
  - Scalar: $$\sigma_{xy}^2 = E[(x - E[x])(y - E[y])]$$
  - Vector: $$\Sigma_{\mathbf{xy}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])^\top]$$

Expectation Exercise

- We know that:

$$\Sigma_{\mathbf{xx}} = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^\top]$$

- Suppose we measure a bunch of samples of $\mathbf{x}$. We compute the first and second moments of $\mathbf{x}$, i.e.,

$$M_{\mathbf{x}} = \sum \mathbf{x}$$

$$M_{\mathbf{xx}} = \sum \mathbf{x}\mathbf{x}^\top$$

- How do we compute Σ using only these moments and the number of samples?

Projecting Covariances

- Suppose I know $\mathbf{x} \sim \mu_{\mathbf{x}}, \Sigma_{\mathbf{x}}$

- How do we handle $\mathbf{y} = \mathrm{A}\mathbf{x} + \mathbf{b}$???

$$\Sigma_{\mathbf{yy}} = E[(\mathbf{y} - E[\mathbf{y}])(\mathbf{y} - E[\mathbf{y}])^\top]$$

- (Algebra) ➔ $\Sigma_{\mathbf{yy}} = \mathrm{A}\Sigma_{\mathbf{xx}}\mathrm{A}^\top$

Properties of the Covariance Matrix

$$\Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

- Symmetric

  $B = C^\top$ why?

- Positive (semi) definite

  $$\mathbf{a}^\top \Sigma \mathbf{a} \geq 0 \quad \text{why?}$$

- Inverse is also positive definite
  - Proof: see next slide

- Determinant → Volume of uncertainty (Product of the Eigenvalues)

Positive (Semi) Definite Properties

1. If $A \geq 0$ and $B \geq 0$ then $A+B \geq 0$

$$\mathbf{x}^\top (A + B)\mathbf{x} = \mathbf{x}^\top A\mathbf{x} + \mathbf{x}^\top B\mathbf{x}$$

2. If either A or B is positive definite, then so is $A+B$; this follows from 1.

3. If $A > 0$, then $A^{-1} > 0$

$$\mathbf{x}^\top A\mathbf{x} = \mathbf{x}^\top A A^{-1} A\mathbf{x} = (A\mathbf{x})^\top A^{-1} (A\mathbf{x}) > 0 \quad \text{if} \quad \mathbf{x} \neq 0$$

4. If $A \geq 0$, then $F^T A F \geq 0$ for any (not necessarily square) matrix F for which $F^T A F$ is defined.

$$\mathbf{x}^\top (F^\top A F)\mathbf{x} = (F\mathbf{x})^\top A (F\mathbf{x}) \geq 0$$

5. If $A > 0$ and F invertible, then $F^T A F > 0$. This follows from 3 and 4.

Correlation Coefficient

- The correlation coefficient is defined as:

$$\rho_{xy} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \qquad\qquad |\rho_{xy}| \leq 1$$

- Covariance matrix in terms of correlation coefficients

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \rho_{n2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix}$$