TU München
Fakultät für Informatik
PD Dr. Rudolph Triebel
John Chiotellis

Machine Learning for Robotics and Computer Vision
Summer term 2016

Homework Assignment 1
Topic 1: Linear Algebra
April 29, 2016

Exercise 1: Warm up

a) What multiple of $a = (1,1,1)$ is closest to the point $b = (2,4,4)$? Find also the closest point to a on the line through b.

There is some vector $p = \lambda a, \lambda \neq 0$ which is closest to b. Then p is perpendicular to the vector $b - p$ which means $p^T(b-p) = 0$. We just need to find λ, so we solve $\lambda a^T(b - \lambda a) = 0$ and get $\lambda = \frac{a^T b}{a^T a}$.
Plugging in the numbers, we get $\lambda = \frac{10}{3}$, so the closest point is $\lambda a = \frac{10}{3}(1,1,1)$.
Equivalently the closest point to a is $\mu b = \frac{10}{36} b = \frac{10}{36}(2,4,4)$.

b) Prove that the trace of $P = aa^T / a^T a$ always equals 1.

We just unfold $aa^T = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} a_1 \dots a_n \end{pmatrix} = \begin{bmatrix} a_1^2 & a_1 a_2 & \dots & a_1 a_n \\ a_2 a_1 & a_2^2 & \dots & a_2 a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n a_1 & a_n a_2 & \dots & a_n^2 \end{bmatrix}$.
Also $a^T a = \sum_i a_i^2$. Therefore the trace of P is $Tr(P) = Tr(aa^T/a^T a) = \frac{a_1^2 + \dots a_n^2}{\sum_i a_i^2} = 1$.

c) Show that the length of Ax equals the length of $A^T x$ if $AA^T = A^T A$.

$||Ax||^2 = (Ax)^T(Ax) = x^T A^T A x = x^T A A^T x = (A^T x)^T (A^T x) = ||A^T x||^2$.

d) Which 2×2 matrix projects the x,y plane onto the line $x + y = 0$?

We are looking for the matrix $A \in \mathbb{R}^{2\times 2}$ that when multiplied with any vector $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ gives us a vector u that is a *projection* of v on the line $x + y = 0$ or otherwise it is a vector $p = \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. This means that $Av = p$ and $p^T(v - p) = 0$.

Solving for $\lambda \neq 0$ we get

$$\begin{aligned}
p^T(v-p) &= 0 \\
\lambda(1 \quad -1)(\begin{pmatrix} x \\ y \end{pmatrix} - \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix}) &= 0 \\
\lambda(x-y) - 2\lambda^2 &= 0 \\
\lambda &= \frac{1}{2}(x-y)
\end{aligned}$$

$$\Rightarrow p = \frac{1}{2}(x-y)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

So we have

$$\begin{aligned}
Av &= p \\
\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2}(x-y)\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
&\Rightarrow \begin{cases} a_{11}x + a_{12}y & = \frac{1}{2}x - \frac{1}{2}y \\ a_{21}x + a_{22}y & = -\frac{1}{2}x + \frac{1}{2}y \end{cases}
\end{aligned}$$

And since we have no other constraint for A, we use the obvious solution

$$A = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Exercise 2: Determinants

a) If a square matrix A has determinant $\frac{1}{2}$, find $\det(2A)$, $\det(-A)$, $\det(A^2)$ and $\det(A^{-1})$.

$$\begin{aligned}
\det(2A) = 2^n \det(A) = 2^n \frac{1}{2} &= 2^{n-1} \\
\det(-A) = (-1)^n \det(A) &= \pm\frac{1}{2} \\
\det(A^2) = \det(AA) = \det(A)\det(A) = (\frac{1}{2})^2 &= \frac{1}{4} \\
\det(A^{-1}) = \det(A)^{-1} = (\frac{1}{2})^{-1} &= 2
\end{aligned}$$

b) Find the determinants of

$$A = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}\begin{bmatrix} 2 & -1 & 2 \end{bmatrix} \quad , \quad U = \begin{bmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 2 \end{bmatrix}, U^T \text{ and } U^{-1}$$

$$\det(A) = 0 \quad \text{(A has rank 1 so it is not invertible)}$$

$$\det(U) = \prod_{\lambda \in \{4,1,2,2\}} \lambda = 16 \quad \text{(product of the eigenvalues which lie on the diagonal on a triangular matrix)}$$

$$\det(U^T) = det(U) = 16$$

$$\det(U^{-1}) = \det(U)^{-1} = \frac{1}{16}$$

Exercise 3: Eigenvalues and Eigenvectors

a) Find the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$, their traces and their determinants.

$$\det(A - \lambda I) = (3 - \lambda)(1 - \lambda)(-\lambda) = 0 \Rightarrow \lambda \in \{3, 1, 0\}$$

To find the eigenvectors we plug in the eigenvalues and solve the linear system $Ax = \lambda x$ for $x \neq 0$. The corresponding eigenvectors are then

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

The trace and determinant are

$$\begin{aligned} \text{Tr}(A) &= 3 + 1 + 0 = 4 \\ \det(A) &= 0 \end{aligned}$$

For matrix B we have

$$\begin{aligned} \det(B - \lambda I) = (-\lambda)(2 - \lambda)(-\lambda) + 2(-2)(2 - \lambda) &= 0 \\ (\lambda^2 - 4)(2 - \lambda) &= 0 \\ (\lambda + 2)(\lambda - 2)(2 - \lambda) &= 0 \\ &\Rightarrow \lambda \in \{-2, 2, 2\} \end{aligned}$$

The corresponding eigenvectors are then

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

The trace and determinant are

$$\mathrm{Tr}(B) = 0 + 2 + 0 = 2$$
$$\det(B) = 2(0 - 4) = -8$$

Typically eigenvectors are normalized to have length 1 but any multiple of an eigenvector is also an eigenvector.

b) Using the characteristic polynomial, find the relationship between the trace, the determinants and the eigenvalues of any square matrix A.

We can factor the characteristic polynomial as a function of λ as

$$\det(A - \lambda I) = p(\lambda) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n) \tag{1}$$

where λ_i are the roots of the polynomial and the eigenvalues of A. We can simply set $\lambda = 0$ and find that

$$\det(A) = p(0) = (-1)^n(-\lambda_1)\cdots(-\lambda_n) = (-1)^n\prod_{i=1}^{n}(-\lambda_i) = (-1)^n\prod_{i=1}^{n}(-1)(\lambda_i)$$
$$= (-1)^n(-1)^n\prod_{i=1}^{n}\lambda_i = (-1)^{2n}\prod_{i=1}^{n}\lambda_i = \prod_{i=1}^{n}\lambda_i$$

So the determinant of a matrix is equal to the product of its eigenvalues.

Let us deal with the trace. Consider the 2×2 case

$$\begin{aligned}
det(A) &= \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \\
det(A - \lambda I) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\
&= (a - \lambda)(d - \lambda) - bc \\
&= ad - bc - \lambda(a + d) + \lambda^2 \\
&= \lambda^2 - \lambda \cdot Tr(A) + \det(A)
\end{aligned}$$

Considering the $n \times n$ case and focusing on the diagonal, we find that

$$det(A - \lambda I) = (-\lambda)^n + (-\lambda)^{n-1} \cdot Tr(A) + \sum_{j=2}^{n-2}\beta_j\lambda^j + \det(A) \tag{2}$$

Comparing equations (1) and (2) we see that

$$Tr(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n = \sum_{i=1}^{n}\lambda_i \tag{3}$$

c) Diagonalize the unitary matrix V to reach $V = U\Lambda U^*$. All $|\lambda| = 1$. $V = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}$

We have

$$\begin{aligned}
\det(V - \lambda I) &= (\frac{1}{\sqrt{3}} - \lambda)(-\frac{1}{\sqrt{3}} - \lambda) - \frac{1}{3}(1+i)(1-i) \\
&= (\frac{1}{\sqrt{3}} - \lambda)(-\frac{1}{\sqrt{3}} - \lambda) - \frac{2}{3} \\
&= -\frac{1}{3} + \lambda^2 - \frac{2}{3} \\
&= \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)
\end{aligned}$$

Eigenvalues are $\lambda \in \{1, -1\}$ and corresponding eigenvectors are

$$x_1 = \frac{1}{\sqrt{1+2c^2}}\begin{pmatrix} 1 \\ c+ic \end{pmatrix} \quad \text{and} \quad x_2 = \frac{1}{\sqrt{1+2c^2}}\begin{pmatrix} -c+ic \\ 1 \end{pmatrix}$$

where $c = \frac{\sqrt{3}-1}{2}$.
Note that we could arrange the eigenvectors differently but since the matrix U is unitary, we have to keep the diagonal entries real. Now we can write matrix U as

$$U = \frac{1}{\sqrt{1+2c^2}}\begin{bmatrix} 1 & -c+ic \\ c+ic & 1 \end{bmatrix}$$

Therefore our decomposition can be written as

$$V = \frac{1}{1+2c^2}\begin{bmatrix} 1 & -c+ic \\ c+ic & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & c-ic \\ -c-ic & 1 \end{bmatrix}$$

d) Suppose T is a 3×3 upper triangular matrix with entries t_{ij}. Compare the entries of T^*T and TT^*. Show that if they are equal, then T must be diagonal. (All normal triangular matrices are diagonal)

Let $T = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$ with $a, b, c, d, e, f \in \mathbb{C}$.

Then

$$\begin{aligned}
T^*T &= \begin{bmatrix} \bar{a} & 0 & 0 \\ \bar{b} & \bar{d} & 0 \\ \bar{c} & \bar{e} & \bar{f} \end{bmatrix}\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} \bar{a}a & \bar{a}b & \bar{a}c \\ \bar{b}a & \bar{b}b + \bar{d}d & \bar{b}c + \bar{d}e \\ \bar{c}a & \bar{c}b + \bar{e}d & \bar{c}c + \bar{e}e + \bar{f}f \end{bmatrix} \\
TT^* &= \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}\begin{bmatrix} \bar{a} & 0 & 0 \\ \bar{b} & \bar{d} & 0 \\ \bar{c} & \bar{e} & \bar{f} \end{bmatrix} = \begin{bmatrix} a\bar{a} + b\bar{b} + c\bar{c} & b\bar{d} + c\bar{e} & c\bar{f} \\ d\bar{b} + e\bar{c} & d\bar{d} + e\bar{e} & e\bar{f} \\ f\bar{c} & f\bar{e} & f\bar{f} \end{bmatrix}
\end{aligned}$$

Now if $TT^* = T^*T$ we see from the diagonal entries that $-b\bar{b} = c\bar{c}$ and $\bar{b}b = e\bar{e}$. So, it must be that $b = c = e = 0$ and therefore T is diagonal.

Exercise 4: Singular Value Decomposition

a) Find the singular values and singular vectors of
$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$

Eigenvalues of $A^T A$ are

$$\begin{aligned} \det(A^T A - \lambda I) &= \lambda(\lambda - 85) = 0 \\ &\Rightarrow \lambda \in \{0, 85\} \end{aligned}$$

Eigenvectors of $A^T A$ are $\begin{pmatrix} -4 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ with norm $\sqrt{17}$.

Eigenvalues of AA^T are also $\lambda \in \{0, 85\}$

Eigenvectors of AA^T are $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ with norm $\sqrt{5}$.

Therefore:

$$A = \frac{1}{\sqrt{17}} \begin{bmatrix} -4 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{85} \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}$$

b) Explain how $U\Sigma V^T$ expresses A as a sum of r rank-1 matrices: $A = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T$

We see the factorization as

$$A = U\Sigma V^T = U(\Sigma V^T) = \begin{bmatrix} u_1 \ldots u_m \end{bmatrix} \left(\begin{bmatrix} \sigma_1 & & 0 & 0 & 0 & 0 \\ & \ddots & & & & 0 \\ 0 & & \sigma_r & 0 & 0 & 0 \\ 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & \ldots & 0 & 0 & \ddots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \right)$$

$$= \begin{bmatrix} u_1 \ldots u_m \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_r v_r^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T + 0 \cdot u_{r+1} + \ldots + 0 \cdot u_m$$

Note that for the rank it holds $r \leq m$ and $r \leq n$.

c) If A changes to $4A$ what is the change in the SVD?

If $A = U\Sigma V^*$ then $4A = 4U\Sigma V^* = U(4\Sigma)V^*$. We apply the scaling to the singular values and leave the singular vectors normalized as they are.

What is the SVD for A^T and for A^{-1} ?

If $A = U\Sigma V^*$ then $A^T = (U\Sigma V^*)^T = V\Sigma^T U^T$ The singular values stay in the diagonal, but the dimensions of matrix Σ swap.

If $A = U\Sigma V^*$ then we can only compute the pseudoinverse $A^+ = (U\Sigma V^*)^+ = (V^*)^{-1}\Sigma^+ U^{-1} = V\Sigma^+ U^*$ Since U, V are unitary, their (conjugate) transpose is also their inverse. The reciprocals of the singular values are in the diagonal.

d) Find the SVD and the pseudoinverse of $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

The SVD of A will be $A = U\Sigma V^*$ where U is 1×1 meaning a scalar and since it is unitary it is 1, therefore $A = \Sigma V^*$.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad AA^T = 4$$

Then

$$\begin{aligned} \det(AA^T - \lambda I) &= 4 - \lambda = 0 \\ &\Rightarrow \lambda = 4 \end{aligned}$$

and

$$\begin{aligned} \det(A^T A - \lambda I) &= \ldots = \lambda^3(\lambda - 4) = 0 \\ &\Rightarrow \lambda \in \{0, 4\} \end{aligned}$$

For $\lambda = 4$ we get one eigenvector $v_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. For $\lambda = 0$ we get three eigenvectors with only one constraint, that the sum of their entries is zero. We choose them to be orthogonal to each other and normalize them, so that matrix V is indeed unitary.

$$v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad v_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \quad v_4 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

AA^T has one eigenvalue $\lambda = 4$, therefore $\sigma = 2$ and $\Sigma = \begin{bmatrix} 2 & 0 & 0 & 0 \end{bmatrix}$ since A has rank 1.

We now can write the SVD of A:

$$A = U\Sigma V^* = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \end{bmatrix} c \begin{bmatrix} c & c & c & c \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ c & c & -c & -c \end{bmatrix}$$

where $c = \frac{1}{\sqrt{2}}$.

The pseudoinverse of A is then

$$A^+ = V\Sigma U^* = c \begin{bmatrix} c & 1 & 0 & c \\ c & -1 & 0 & c \\ c & 0 & 1 & -c \\ c & 0 & -1 & -c \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \frac{c^2}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

For B we have

$$B = U\Sigma V^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and therefore pseudoinverse

$$B^+ = V\Sigma^+ U^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Finally, for C we have

$$C = U\Sigma V^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and therefore pseudoinverse

$$C^+ = V\Sigma^+ U^* = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$

Topic 2: Probabilistic Reasoning

Exercise 5: Follow the robot

Assume you get your hands on a robot that can measure its distance to a wall in front of it. You model this using a continuous random variable with a Normal distribution $\mathcal{N}(x; \mu, \sigma^2)$.

a) The robot also has a camera on board that is not color-calibrated correctly so the color mapping is probabilistic and looks like the following table:

z \ x	**R**	**G**	**B**
R	0.8	0.1	0.1
G	0.1	0.6	0.2
B	0.1	0.3	0.7

For instance, the probability that the robot reads blue while the true color is green is $p(z = B|x = G) = 0.3$

Assume the robot is located in a white room with 5 boxes: 2 red, 2 green and a blue one. The robot moves towards a box and the camera reads green. How likely is it that the box is actually green?

We simply apply Bayes rule:

$$\begin{aligned} p(X = G|Z = G) &= \frac{p(Z = G|X = G)p(X = G)}{p(Z = G)} \\ &= \frac{p(Z = G|X = G)p(X = G)}{\sum_{x \in \{R,G,B\}} p(Z = G|X = x)p(X = x)} \\ &= \frac{0.6\frac{2}{5}}{0.1\frac{2}{5} + 0.6\frac{2}{5} + 0.2\frac{1}{5}} \\ &= \frac{0.24}{0.04 + 0.24 + 0.04} \\ &= \frac{0.24}{0.32} = \frac{3}{4} = 0.75 \end{aligned}$$

b) The robot's distance sensor can be modeled using a continuous random variable with a Normal distribution with $\sigma_1 = 0.3$ m. Express the sensor model $p(z|x)$ in the full form (not the shorthand notation).

$$p(z \mid x) = \mathcal{N}(z|x, \sigma_1^2) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\frac{(z-x)^2}{\sigma_1^2}} = \frac{1}{0.3\sqrt{2\pi}} e^{-5.55(z-x)^2} \tag{4}$$

where $\sigma_1 = 0.3$ is the sensor noise.

c) Now the robot moves into another room that is empty. Initially it knows it is located at the door (x=0). The robot can execute *move* commands but the result of the action is not always perfect. Assume that the robot moves with constant speed v. The motion can also be modeled with a Gaussian with deviation $\sigma_2 = 0.1$ m. Write the motion model $p(x_t|x_{t-1}, u_t)$.

The motion can also be modeled with a Gaussian. We just need to think about what is our mean and what is our variance. The variance is given as the actuator noise $\sigma_2 = 0.1$. Our mean is the position we expect our robot to be at, after the motion u_t. Since our robot moves with constant speed v, the expected position is simply $\mu = x_{t-1} + v\Delta t$. Therefore we have

$$\begin{aligned} p(x_t|x_{t-1}, u_t) &= \mathcal{N}(x_t|x_{t-1} + v\Delta t, \sigma_2^2) \\ &= \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2}\frac{(x_t-(x_{t-1}+v\Delta t))^2}{\sigma_2^2}} \\ &= \frac{10}{\sqrt{2\pi}} e^{-50(x_t-(x_{t-1}+v\Delta t))^2} \end{aligned}$$

d) You let the robot run in the room with a speed of 1 m/s. The robot only runs forward and it updates its belief every second. Assume you get the following sensor measurements in the first 3 seconds: $\{z_1 = 1.2, z_2 = 1.6, z_3 = 2.5\}$.

Further assume that the position can only take discrete values from 0 to 5. Where does the robot believe it is located with respect to the door after 3 seconds? How certain is it about its location?

We model the state variable x as a discrete random variable with values between 0 and 5, where 0 means that the robot is at the door. We want to compute the robot's belief. Initially, the robot knows it is located at the door (x=0), therefore we have $Bel(x_0 = 0) = 1$. We then use the Bayes filter algorithm to compute the belief after 3 seconds, namely $Bel(x_3)$. Since it is a recursive algorithm we have to compute the belief at every time step. The general equation of the Bayes filter is:

$$Bel(x_t) = \eta \quad p(z_t|x_t) \int p(x_t|u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1} \tag{5}$$

Our u_i, namely our action is always the same: move with constant speed $v = 1m/s$. Here is the first step:

$$\begin{aligned} Bel(x_1) &= \eta_1 p(z_1|x_1) \int p(x_1|u_1, x_0) Bel(x_0) dx_0 && (6) \\ &= \eta_1 p(z_1|x_1) \int p(x_1|u_1, x_0) dx_0 && (7) \\ &= \eta_1 \mathcal{N}(z_1|x_1, \sigma_1^2) \sum_{x_0=0}^{5} \mathcal{N}(x_1|x_0 + 1, \sigma_2^2) && (8) \\ &= \eta_1 \mathcal{N}(z_1|x_1, \sigma_1^2) \mathcal{N}(x_1|x_0 + 1, \sigma_2^2) && (9) \end{aligned}$$

Let us separate the computation to the *motion* and the *sensing* part. Since we have $Bel(x_0 = 0) = 1$ we begin from the motion u_1.

$$p(x_1|u_1, x_0) = \mathcal{N}(x_1|x_0 + 1, \sigma_2^2) \tag{10}$$

The belief of the robot after the first motion can be estimated as

$$Bel'(x_1) = \sum_{x_0=0}^{5} \mathcal{N}(x_1|x_0 + 1, \sigma_2^2) \tag{11}$$

Now we take into account the sensor measurement z_1:

$$p(z_1|x_1) = \mathcal{N}(z_1|x_1, \sigma_1^2) \tag{12}$$

Therefore

$$Bel(x_1) = \eta_1 p(z_1|x_1) Bel'(x_1) \tag{13}$$

Since our positions are restricted to a space $x_t \in \{0, 1, 2, 3, 4, 5\}$, we can compute our normalizers η_i using (the inverse of) the sum of the probabilities for all possible states.

$$\eta_1^{-1} = \sum_{x_1'=0}^{5} Bel(x_1 = x_1') \tag{14}$$

If we recursively substitute the beliefs we get:

$$\begin{aligned} Bel(x_3) &= \eta_3 p(z_3|x_3) \int p(x_3|u_3, x_2) Bel(x_2) dx_2 \\ &= \eta_3 p(z_3|x_3) \int p(x_3|u_3, x_2) \eta_2 p(z_2|x_2) \int p(x_2|u_2, x_1) Bel(x_1) dx_1 dx_2 \end{aligned}$$

Plugging the numbers in we get the following table:

x	0	1	2	3	4	5
$Bel(x_0)$	1	0	0	0	0	0
$Bel(x_1)$	0.0001	0.9998	6.8798e-24	2.6318e-95	5.5977e-215	0
$Bel(x_2)$	3.9381e-11	0.0762	0.9237	8.3773e-27	3.7763e-59	2.7476e-138
$Bel(x_3)$	2.6757e-26	2.3196e-07	0.2499	0.7500	2.1622e-27	4.3797e-63

We can see that the robot believes that it is $3m$ away from the door and is about 75% certain.

Exercise 6: An overview of ML methods

Try to find (for example by internet search or from the book (Bishop or)) at least 5 examples for learning techniques that have not been discussed in class. Describe these techniques briefly and classify them with respect to the categories presented in the lecture.

Here are some examples of learning algorithms:

- Mean-shift clustering: Unsupervised learning
- Perceptron algorithm: Discriminant function
- Neural Networks: Discriminative model
- Bayes classifier: Generative model
- Conditional Random Field: Discriminative model
- AdaBoost: Discriminant function

For a detailed explanation, please see the textbook *Pattern Recognition and Machine Learning* by *C.M. Bishop* or the slides.

The next exercise class will take place on **May 13th, 2016**.

For downloads of slides and of homework assignments and for further information on the course see

`https://vision.in.tum.de/teaching/ss2016/mlcv16`
