PROBABILISTIC ROBOTICS:
MOBILE ROBOT LOCALIZATION: GRID AND MONTE CARLO

Pierre-Paul TACHER

1

Let d the dimension of state space. For simplicity we assume the dimension d_1 of the measurement space satisfy $d_1 = O(d)$.

The EKF algorithm implies the computation of Jacobians and evaluations of functions whose complexities depend on the particular model. Thus we will ignore it here (we assume complexity is $O(1)$). The time complexity of EKF localization is dominated by the constant number of order d matrix multiplications and inversions; it can be shown [1] the time complexity of such operation is the same, and current state of the art [2] guarantees complexity $O(d^\omega)$, where $\omega < 2.3728639$. Thus this is also the time complexity of EKF algorithm. The memory complexity is $O(d^2)$.

Let's consider the grid algorithm; for now we use a common discretization resolution h for all axes and not dependant on d so that the number of grid cells is $O(h^d)$. The loop over the grid cells contains $O(h^d)$ additions / multiplications, hence a time complexity of $O(h^{2d})$. The memory complexity is $O(h^d)$.

The time or memory complexity of MCL is $O(M)$, where M is the number of particles. This number have to be scaled with d to represent accurately the probability distribution. We will roughly estimate this dependency: Let M_d an estimate of the number of particles needed to represent gaussian r.v. $\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} \hookrightarrow \mathcal{N}(0, I_d)$. In dimension $d+1$, let's estimate how many particles are needed to represent gaussian r.v. $\begin{bmatrix} X_1 \\ \vdots \\ X_d \\ X_{d+1} \end{bmatrix} \hookrightarrow \mathcal{N}(0, I_{d+1})$; for a fixed $x_{d+1} \in \mathbb{R}$, the d dimensional gaussian r.v. $\begin{bmatrix} X_1 \\ \vdots \\ X_d \\ x_{d+1} \end{bmatrix}$ conditioned on $(X_{d+1} = x_{d+1})$ needs approximately M_d particles with n+1$^{\text{th}}$ coordinate x_{d+1} to be represented. And we need to choose at least M_1 values of x_{d+1} to hope to represent the gaussian marginals $\begin{bmatrix} x_1 \\ \vdots \\ x_d \\ X_{d+1} \end{bmatrix}$. So we have $M_{d+1} \approx M_d \times M_1$ so that we end in exponential dependency $M_d \approx M_1^d$.

[1] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford: *Introduction to Algorithms (3rd ed.)*, MIT Press and McGraw-Hill, Theorems 28.1 & 28.2(2009)

[2] Le Gall, François: *Powers of tensors and fast matrix multiplication*, Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation (2014)

2

The key point is that the r.v. $Z_1, Z_2, \ldots, Z_k$ conditioned on x_t are assumed to be independant. The loop over the k measurements in EKF localization algorithm can be interpreted as simple iterations of the basic EKF where there is no motion between the measurement updates. We have,

$$p(x_t \mid z_1) = \eta p(z_1 \mid x_t) \times p(x_t)$$

and

$$\begin{aligned} \forall i \in [\![2,k]\!], \quad p(x_t \mid z_{1:i+1}) &= \eta p(z_{i+1} \mid x_t, z_{1:i}) \times p(x_t \mid z_{1:i}) \\ &= \eta p(z_{i+1} \mid x_t) \times p(x_t \mid z_{1:i}) \end{aligned}$$

A straighforward induction shows that $\forall i \in [\![1,k]\!]$ the law of x_t conditioned on $z_{1:i}$ is gaussian; since it is known that the law of Z_i conditioned on x_t is gaussian $\hookrightarrow \mathcal{N}(h(\bar{\mu}, j(i)) + H_{j(i)}(x_t - \bar{\mu}), Q_t)$,the mean and covariance matrices are computed by the relations already derived in the original Kalmann filter

$$\begin{aligned} \forall i \in [\![1,k]\!], \quad K_i &= \Sigma_{i-1} H_{j(i)}^T (H_{j(i)} \Sigma_{i-1} H_{j(i)}^T + Q)^{-1} \\ \mu_i &= \mu_{i-1} + K_i(z_i - h(\bar{\mu}, j(i))) \\ \Sigma_i &= (1 - K_i H_{j(i)}) \Sigma_{i-1} \end{aligned}$$

where $\mu_0 = \bar{\mu}$, $\Sigma_0 = \bar{\Sigma}$, $\mu_k = \mu$ and $\Sigma_k = \Sigma$. These are precisely the updates in the loop over measurements in the EKF localization algorithm.

3

Each $w_t^{[m]} = p(z_t \mid x_t^{[m]})$ can be seen as independant realization of the random variable $p(z_t \mid X_t)$, where X_t is the random variable giving the position of a particule at time t, conditioned on past actions and observations $z_{1:t-1}, u_{1:t}$. This r.v. has a finite expectancy given by

$$\begin{aligned} E(p(z_t \mid X_t)) &= \int_\Omega p(z_t \mid X_t) \,\mathrm{d}P \\ &= \int_\Omega p(z_t \mid X_t, z_{1:t-1}, u_{1:t}) \,\mathrm{d}P \\ &= \int_\Omega \frac{p(X_t \mid z_{1:t}, u_{1:t}) p(z_t \mid z_{1:t-1}, u_{1:t})}{p(X_t \mid z_{1:t-1}, u_{1:t})} \,\mathrm{d}P \\ &= p(z_t \mid z_{1:t-1}, u_{1:t}) \int_\mathbb{R} \frac{p(x_t \mid z_{1:t}, u_{1:t})}{p(x_t \mid z_{1:t-1}, u_{1:t})} p(x_t \mid z_{1:t-1}, u_{1:t}) \,\mathrm{d}x_t \\ &= p(z_t \mid z_{1:t-1}, u_{1:t}) \underbrace{\int_\mathbb{R} p(x_t \mid z_{1:t}, u_{1:t}) \,\mathrm{d}x_t}_{=1} \\ &= p(z_t \mid z_{1:t-1}, u_{1:t}) \end{aligned}$$

where we assumed $p(X_t \mid z_{1:t-1}, u_{1:t}) > 0$ almost surely. The strong law of large numbers[3] shows that, almost surely,

$$\lim_{M \to +\infty} \frac{1}{M} \sum_{m=0}^{M} w_t^{[m]} = p(z_t \mid z_{1:t-1}, u_{1:t})$$

[3]Durrett, Rick: *Probability: Theory and Examples (4th ed.)*, Cambridge University Press, Theorems 2.4.1(2013)

4

4.1. Let us define the r.v. $Y \in \mathbb{R}^4$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix}$$

where $Y_i \in [\![0, N]\!]$ is the r.v. giving the number of particles in location x_i. We have $\Sigma_{i=1}^4 Y_i = N$. Let us define the set

$$D_N = \{y \in [\![0, N]\!]^4, \quad \sum_{i=1}^{4} y_i = N\}$$

Each of the location being equiprobable, Y follows a multinomial law $Y \hookrightarrow \mathcal{M}(N, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ (cf. A.1 for some details on this distribution). Thus,

$$\begin{aligned} \forall y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \in D_N, \quad P(Y = y) &= \frac{N!}{y_1!y_2!y_3!y_4!} (\frac{1}{4})^{y_1} (\frac{1}{4})^{y_2} (\frac{1}{4})^{y_3} (\frac{1}{4})^{y_4} \\ &= \frac{N!}{y_1!y_2!y_3!y_4!} (\frac{1}{4})^N \end{aligned}$$

Let $X_s^N \in \{x_1, x_2, x_3, x_4\}$ the r.v. which gives the position of the particle which is resampled by the algorithm. X is the r.v. of the position of a particle before resampling $X \hookrightarrow \mathcal{U}(\{x_1, x_2, x_3, x_4\})$.

As usual we use $P(X_s^N = x_i \mid z)$ as short of $P(X_s^N = x_i \mid Z = 1)$ and $P(X_s^N = x_i \mid \neg z)$ as short of $P(X_s^N = x_i \mid Z = 0)$. The particle is resampled according to the normalized weights $w^{[n]} = p(z \mid x), n \in [\![1, N]\!]$ and the initial sampling of the N particles is assumed to be independant of Z, so that

$$\begin{aligned} P(X_s^N = x_i \mid z) &= \sum_{y \in D_N} P(X = x_i \cap Y = y \mid z) \\ &= \sum_{y \in D_N} P(X_s^N = x_i \mid Y = y, z) \times \underbrace{P(Y = y \mid z)}_{P(Y=y)} \\ &= \frac{N!}{4^N} \sum_{y \in D_N} \frac{1}{y_1!y_2!y_3!y_4!} \frac{y_i \times p(z \mid x_i)}{\sum_{j=1}^4 y_j \times p(z \mid x_j)} \end{aligned}$$

and similarly

$$P(X_s^N = x_i \mid \neg z) = \frac{N!}{4^N} \sum_{y \in D_N} \frac{1}{y_1!y_2!y_3!y_4!} \frac{y_i \times p(\neg z \mid x_i)}{\sum_{j=1}^4 y_j \times p(\neg z \mid x_j)}$$

For some values of the number of particles, we compute the distributions with Matlab function `post_resampling_dist.m`. The results are:

i	1	2	3	4
$P(X_s^1 = x_i \mid z)$	0.25	0.25	0.25	0.25
$P(X_s^1 = x_i \mid \neg z)$	0.25	0.25	0.25	0.25
$P(X_s^2 = x_i \mid z)$	0.3681	0.3042	0.1639	0.1639
$P(X_s^2 = x_i \mid \neg z)$	0.1392	0.2563	0.3023	0.3023
$P(X_s^3 = x_i \mid z)$	0.4302	0.3145	0.1277	0.1277
$P(X_s^3 = x_i \mid \neg z)$	0.1107	0.2506	0.3194	0.3194
$P(X_s^4 = x_i \mid z)$	0.4661	0.3141	0.1099	0.1099
$P(X_s^4 = x_i \mid \neg z)$	0.0966	0.2462	0.3271	0.3271
$P(X_s^5 = x_i \mid z)$	0.4886	0.3115	0.1000	0.1000
$P(X_s^5 = x_i \mid \neg z)$	0.0940	0.2433	0.3314	0.3314
$P(X_s^6 = x_i \mid z)$	0.5037	0.3086	0.0939	0.0939
$P(X_s^6 = x_i \mid \neg z)$	0.0905	0.2413	0.3341	0.3341
$P(X_s^7 = x_i \mid z)$	0.5143	0.3060	0.0898	0.0898
$P(X_s^7 = x_i \mid \neg z)$	0.0833	0.2398	0.3360	0.3360
$P(X_s^8 = x_i \mid z)$	0.5221	0.3039	0.0870	0.0870
$P(X_s^8 = x_i \mid \neg z)$	0.0867	0.2387	0.3373	0.3373
$P(X_s^9 = x_i \mid z)$	0.5281	0.3021	0.0849	0.0849
$P(X_s^9 = x_i \mid \neg z)$	0.0854	0.2378	0.3384	0.3384
$P(X_s^{10} = x_i \mid z)$	0.5328	0.3006	0.0833	0.0833
$P(X_s^{10} = x_i \mid \neg z)$	0.0845	0.2371	0.3392	0.3392
$P(X_s^{100} = x_i \mid z)$	0.5679	0.2872	0.0724	0.0724
$P(X_s^{100} = x_i \mid \neg z)$	0.0776	0.2314	0.3455	0.3455
$P(X_s^{200} = x_i \mid z)$	0.5697	0.2865	0.0719	0.0719
$P(X_s^{200} = x_i \mid \neg z)$	0.0773	0.2311	0.3458	0.3458
$P(X_s^{300} = x_i \mid z)$	0.5703	0.2862	0.0718	0.0718
$P(X_s^{300} = x_i \mid \neg z)$	0.0772	0.2310	0.3459	0.3459

When the number of particles goes to infinity, we write

$$\frac{Y_i \times p(z \mid x_i)}{\sum_{j=1}^{4} Y_j \times p(z \mid x_j)} = \frac{\frac{Y_i}{N} \times p(z \mid x_i)}{\sum_{j=1}^{4} \frac{Y_j}{N} \times p(z \mid x_j)}$$

The observed frequencies goes to theoric probabilities, that is

$$\lim_{N \to +\infty} \frac{Y_i}{N} = P(X = X_i) = \frac{1}{4} \qquad \text{a.s.}$$

(observe that $Y_i = \sum_{n=1}^{N} 1_{X^{[n]}=x_i}$ and apply the law of large numbers to indicator r.v. $1_{X=x_i}$), we understand intuitively the limit distribution

$$\begin{aligned}
\lim_{N\to+\infty} P(X_s^N = x_i \mid z) &= \frac{\frac{1}{4} \times p(z \mid x_i)}{\sum_{j=1}^{4} \frac{1}{4} \times p(z \mid x_j)} \\
&= \frac{P(X = x_i) \times p(z \mid x_i)}{\sum_{j=1}^{4} P(X = x_j) \times p(z \mid x_j)} \\
&= P(X = x_i \mid z) \\
\lim_{N\to+\infty} P(X_s^N = x_i \mid \neg z) &= P(X = x_i \mid \neg z)
\end{aligned}$$

which is precisely the target distribution we want to sample from. A more rigorous derivation of the last fact can be found in A.2.

i	1	2	3	4
$P(X = x_i \mid z)$	0.5714	0.2857	0.0714	0.0714
$P(X = x_i \mid \neg z)$	0.0769	0.2308	0.3462	0.3462

4.2. We note KL_i and $\mathrm{KL}_i^\neg$ the divergence from distributions of X_s^i conditioned on $(Z=1)$ and $(Z=0)$ respectively, to target distribution. I use base 2 log.

KL_1	0.4784
$\mathrm{KL}_1^\neg$	0.1676
KL_2	0.1657
$\mathrm{KL}_2^\neg$	0.0347
KL_3	0.0748
$\mathrm{KL}_3^\neg$	0.0127
KL_4	0.0401
$\mathrm{KL}_4^\neg$	0.0063
KL_5	0.0242
$\mathrm{KL}_5^\neg$	0.0038
KL_6	0.0160
$\mathrm{KL}_6^\neg$	0.0025
KL_7	0.0113
$\mathrm{KL}_7^\neg$	0.0018
KL_8	0.0083
$\mathrm{KL}_8^\neg$	0.0013
KL_9	0.0064
$\mathrm{KL}_9^\neg$	0.0010
KL_{10}	0.0051
$\mathrm{KL}_{10}^\neg$	8.3014×10^{-4}
KL_{100}	4.1949×10^{-5}
$\mathrm{KL}_{100}^\neg$	7.4765×10^{-6}
KL_{200}	1.0380×10^{-5}
$\mathrm{KL}_{200}^\neg$	1.8589×10^{-6}
KL_{300}	4.5978×10^{-6}
$\mathrm{KL}_{300}^\neg$	8.2467×10^{-7}

4.3. In the trivial case when $\forall i \in [\![1,4]\!], p(z \mid x_i) = \frac{1}{4}$, we have

$$\begin{aligned}\forall N \in \mathbb{N}^*, \forall i \in [\![1,4]\!], \quad P(X_s^N = x_i \mid z) &= P(X = x_i \mid z) \\ &= \frac{1}{4}\end{aligned}$$

5

To do.

6

See file `histogram_filter_3.m`.

Appendix A

In this appendix, I suppose known by the reader what it means to integrate with respect to a measure, as well as some basic integration properties.

A.1.

Definition. Let $(\Omega, \mathcal{F}, P)$ a probability space.Let $k \in \mathbb{N}^*$ fixed. For each $n \in \mathbb{N}^*$, we consider a partition $(\mathcal{A}_i^n)_{1 \leqslant i \leqslant k}$of Ω, where $\mathcal{A}_i^n \in \mathcal{F}$. We assume the collection of subsets indexed on n are independant and

$$\forall n \in \mathbb{N}^*, \forall i \in [\![1, k]\!], \quad P(A_i^n) = p_i$$

with $p_i > 0$ and $\sum_{i=1}^n p_i = 1$. Let 's define two $\mathbb{R}^k$ valued r.v.

$$X^n = \begin{bmatrix} 1_{\mathcal{A}_1^n} \\ \vdots \\ 1_{\mathcal{A}_k^n} \end{bmatrix}$$

$$Y^n = \sum_{j=1}^n X^j$$

We say that Y^n follows the multinomial law with parameters n, p_1,. . . ,p_{k-1} which is written $Y^n \hookrightarrow \mathcal{M}(n, p_1, \ldots, p_{k-1})$($p_k$ is determined by $\sum_{j=1}^k p_k = 1$).

Proposition. Let $D_n = \{y \in [\![0, n]\!]^k, \quad \sum_{i=1}^k y_i = n\}$; then,

$$\forall y \in D_n, \quad P(Y^n = y) = \frac{n!}{y_1!y_2! \ldots y_k!} p_1^{y_1} p_2^{y_2} \ldots p_k^{y_k}$$

Proof. Let $(e_j)_{j \in [\![1,k]\!]}$ be the canonical basis of $\mathbb{R}^k$.

$$\begin{aligned} P(X^n = e_j) &= P(A_j^n) \\ &= p_j \end{aligned}$$

Since

$$\begin{aligned} \sum_{i=1}^k Y_i^n &= \sum_{i=1}^k \sum_{j=1}^n X_i^j \\ &= \sum_{j=1}^n \sum_{i=1}^k 1_{A_i^j} \\ &= \sum_{j=1}^n 1_\Omega \\ &= n \end{aligned}$$

we have $Y^n(\Omega) \subset D_n$. Let now $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \in D_n$; We note $\mathcal{P}([\![1, n]\!])^{y_1, y_2, \ldots, y_k}$ the set of partitions of $[\![1, n]\!]$ containing exactly k subsets J_i satisfying the conditions $|J_1| = y_1$,$|J_2| = y_2$,. . . ,$|J_k| = y_k$. The number of such partitions is

$$\begin{aligned} |\mathcal{P}([\![1, n]\!])^{y_1, y_2, \ldots, y_k}| &= \binom{n}{y_1}\binom{n - y_1}{y_2} \cdots \binom{n - (y_1 + y_2 + \cdots + y_{k-2})}{y_{k-1}} \\ &= \frac{n!}{y_1!y_2! \ldots y_k!} \end{aligned}$$

We observe that there are as many ways to realize the event $(Y^n = y)$ as there are to choose such a partition, and each of it has the same probability; formally, we write the event as a disjoint union:

$$(Y^n = y) = \biguplus_{\{J_1,\dots,J_k\}\in\mathcal{P}([\![1,n]\!])^{y_1,y_2,\dots,y_k}} \left[\bigcap_{j_1\in J_1}(X^{j_1}=e_1)\cap\bigcap_{j_1\in J_1}(X^{j_2}=e_2)\cap\cdots\cap\bigcap_{j_k\in J_k}(X^{j_k}=e_k)\right]$$

and using the independance of r.v. X^j:

$$\begin{aligned}P(Y^n = y) &= \sum_{\{J_1,\dots,J_k\}\in\mathcal{P}([\![1,n]\!])^{y_1,y_2,\dots,y_k}} p_1^{y_1}p_2^{y_2}\dots p_k^{y_k}\\ &= \frac{n!}{y_1!y_2!\dots y_k!}p_1^{y_1}p_2^{y_2}\dots p_k^{y_k}\end{aligned}$$

□

A.2. In this part we prove the result $\lim_{N\to+\infty} P(X_s^N = x_i \mid z) = P(X \mid z)$ which theoretically justifies the particle filter algorithm. We need the following theorem [4]:

Definition. We say that sequence of r.v. X_n converges **almost surely** to r.v. X if

$$P(\lim_{n\to+\infty} X_n = X) = 1$$

Definition. We say that sequence of r.v. X_n converges **in probability** to r.v. X if

$$\forall\epsilon > 0, \quad \lim_{n\to+\infty} P(|X_n - X| > \epsilon) = 0$$

Remark. Almost sure convergence implies convergence in probability.

Theorem. *Bounded convergence theorem.* Let $(X_n)_{n\in\mathbb{N}}$ a sequence of r.v.s with $\forall n, X_n \leqslant M$, and $X_n \to X$ in probability. Then,

$$\lim_{n\to+\infty} E[X_n] = E[X]$$

Proof. Let $\epsilon > 0$, $G_n = \{\omega\in\Omega, |X_n(\omega) - X(\omega)| < \epsilon\}$ and $B_n = \Omega - G_n$.

$$\begin{aligned}|E[X_n] - E[X]| &= \left|\int X_n\,\mathrm{d}P - \int X\,\mathrm{d}P\right|\\ &= \left|\int (X_n - X)\,\mathrm{d}P\right|\\ &\leqslant \int |X_n - X|\,\mathrm{d}P\\ &= \int_{G_n}|X_n - X|\,\mathrm{d}P + \int_{B_n}|X_n - X|\,\mathrm{d}P\\ &\leqslant \epsilon P(G_n) + 2MP(B_n)\\ &\leqslant \epsilon + 2MP(B_n)\end{aligned}$$

Convergence in probability implies that $P(B_n) \to 0$ and since ϵ can be chosen arbitrarily this completes the proof. □

[4]Durrett, Rick: *Probability: Theory and Examples (4th ed.)*, Cambridge University Press, Theorems 1.5.3(2013)

We go back to our problem :

$$\begin{aligned}
P(X_s^N = x_i \mid z) &= \int_{\mathbb{R}} 1_{x=x_i} \, \mathrm{d}P_{X_s^N}^{Z=z}(x) \\
&= \int_{\mathbb{R}^4} (\int_{\mathbb{R}} 1_{x=x_i} \, \mathrm{d}P_{X_s^N}^{Y=y,Z=z}(x)) \, \mathrm{d}P_Y(y) \\
&= \int_{\mathbb{R}^4} P(X_s^N = x_i \mid Y = y, Z = z) \, \mathrm{d}P_Y(y) \\
&= \int_{\mathbb{R}^4} \frac{y_i p(z \mid x_i)}{y_1 p(z \mid x_1) + y_2 p(z \mid x_2) + y_3 p(z \mid x_3) + y_4 p(z \mid x_4)} \, \mathrm{d}P_Y(y) \\
&= \int \frac{Y_i p(z \mid x_i)}{\sum_{j=1}^4 Y_j p(z \mid x_j)} \, \mathrm{d}P \\
&= E[F^N]
\end{aligned}$$

where F^N is the r.v.

$$\begin{aligned}
F^N &= \frac{Y_i p(z \mid x_i)}{\sum_{j=1}^4 Y_j p(z \mid x_j)} \\
&= \frac{\frac{Y_i}{N} p(z \mid x_i)}{\sum_{j=1}^4 \frac{Y_j}{N} p(z \mid x_j)}
\end{aligned}$$

We have seen this r.v. converges in probability to $\frac{p(z|x_i)}{\sum_{j=1}^4 p(z|x_j)}$, and since it bounded by 1,

$$\begin{aligned}
P(X_s^N = x_i \mid z) &\to E[\frac{p(z \mid x_i)}{\sum_{j=1}^4 p(z \mid x_j)}] \\
&= \frac{p(z \mid x_i)}{\sum_{j=1}^4 p(z \mid x_j)} \\
&= P(X = x_i \mid z)
\end{aligned}$$