Technion – Israel Institute of Technology

HW2

Vision Aided Navigation

086761

Alon Spinner	305184335	alonspinner@gmail.com
Sher Hazan	308026467	sherhazan@campus.technion.ac.il

November 24, 2021

Basic Probability and Bayesian Inference

Question 1 : Consider a random vector $x \in \mathbb{R}^n$ with a Gaussian distribution, written in covariance form $x \sim N(\mu, \Sigma)$. Show the corresponding information form $x \sim N^{-1}(\eta, \Lambda)$ is:

$$p(x) = N^{-1}(\eta, \Lambda) = \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}x^T \Lambda x + \eta^T x\right)$$

$$\Lambda = \Sigma^{-1}, \eta = \Lambda\mu \rightarrow \Sigma = \Lambda^{-1}, \mu = \Lambda^{-1}\eta$$

$$p(x) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) =$$

$$= \frac{1}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x-\Lambda^{-1}\eta)^T \Lambda (x-\Lambda^{-1}\eta)\right) =$$

$$= \frac{1}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x-\Lambda^{-1}\eta)^T \Lambda (x-\Lambda^{-1}\eta)\right) =$$

$$= \frac{1}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x^T\Lambda x - x^T\Lambda\Lambda^{-1}\eta - (\Lambda^{-1}\eta)^T\Lambda x + (\Lambda^{-1}\eta)^T\Lambda\Lambda^{-1}\eta)\right) =$$

$$= \frac{1}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x^T\Lambda x - x^T\eta - \eta^T\Lambda^{-T}\Lambda x + \eta^T\Lambda^{-T}\eta)\right) =$$

$$= \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-T}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x^T\Lambda x - x^T\eta - \eta^T\Lambda^{-T}\Lambda x)\right) =$$

since Σ is symmetric, Σ^{-1} is symmetric too, $\Sigma^{-1} = \Lambda$ so Λ, Λ^{-1} are symmetric.

$$= \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x^T\Lambda x - x^T\eta - \eta^T\Lambda^{-1}\Lambda x)\right) =$$

$$= \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}(x^T\Lambda x - x^T\eta - \eta^T x)\right) =$$

$$= \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}x^T\Lambda x + \frac{1}{2}x^T\eta + \frac{1}{2}\eta^T x\right) =$$

$x^T\eta$ is scalar, there for it is equal to its transpose $(x^T\eta)^T = \eta^T x$.

$$= \boxed{\frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1}\eta\right)}{\sqrt{\det(2\pi\Lambda^{-1})}} \exp\left(-\frac{1}{2}x^T\Lambda x + \eta^T x\right)}$$

Question 2 : Consider a standard observation model involving a random variable $x \in \mathbb{R}^n$.

$$z = h(x) + v \,, v \sim N(0, \Sigma_v)$$

and assume the initial belief regarding the state x is a Gaussian with mean $\hat{x}_0$ and covariance Σ_0.

(a) Write an expression for the prior $p(x)$ and the measurement likelihood $p(z|x)$.

$$p(x) = \boxed{\frac{1}{\sqrt{\det(2\pi\Sigma_0)}} \exp\left(-\frac{1}{2}(x - \hat{x}_0)^T \Sigma_0^{-1}(x - \hat{x}_0)\right)}$$

$$p(z|x) = \boxed{\frac{1}{\sqrt{(det(2\pi\Sigma_v))}} exp\left(-\frac{1}{2}\big(z - h(x)\big)^T {\Sigma_v}^{-1}\big(z - h(x)\big)\right)}$$

(b) A measurement z_1 is acquired. Assuming the measurement was generated by the measurement model (1), write an expression for the posterior probability $p\ (x|z_1)$ in terms of $p(x)$ and the measurement likelihood $p(z|x)$.

$$p(x|z_1) \underset{BR}{=} \frac{p(x)p(z_1|x)}{p(z_1)} = \boxed{\eta p(x)p(z_1|x)}$$

(c) Derive expressions for the posteriori mean $\hat{x}_1$ and covariance Σ_1 such that $p\ (x|z_1) =\ N\ (\hat{x}_1, \Sigma_1)$.

$$\hat{x}_1 = argmax\big(p(x|z_1)\big) = argmin\big(-\log\big(\eta p(x)p(z_1|x)\big)\big)$$

$$Plugging\ in\ values, we\ notice\ that:$$

$$\eta p(x)p(z_1|x) = \eta \frac{\exp\left(-\frac{1}{2}(x - \hat{x}_0)^T \Sigma_0^{-1}(x - \hat{x}_0)\right) exp\left(-\frac{1}{2}\big(z - h(x)\big)^T {\Sigma_v}^{-1}\big(z - h(x)\big)\right)}{\sqrt{\det(2\pi\Sigma_0)}\sqrt{(det(2\pi\Sigma_v))}}$$

$$-\log\big(\eta p(x)p(z_1|x)\big) = -Const\left(\left(-\frac{1}{2}(x - \hat{x}_0)^T \Sigma_0^{-1}(x - \hat{x}_0)\right) + \left(-\frac{1}{2}\big(z - h(x)\big)^T {\Sigma_v}^{-1}\big(z - h(x)\big)\right)\right)$$

$$where\ Const > 0$$

$$We\ define:$$

$$J(x,z) = -\frac{1}{2}||x - \hat{x}_0||^2_{\Sigma_0} - \frac{1}{2}||z - h(x)||^2_{\Sigma_v}$$

$$such\ that:$$

$$-\log\big(\eta p(x)p(z_1|x)\big) = \ \mathrm{Const} \cdot \left(||x - \hat{x}_0||^2_{\Sigma_0} + ||z - h(x)||^2_{\Sigma_v}\right) = Const \cdot J(x,z)$$

$$\hat{x}_1 =\ argmin\big(-\log\big(\eta p(x)p(z_1|x)\big)\big) = argmin\big(Const \cdot J(x,z)\big) = argmin\big(J(x,z)\big)$$

$$since\ h(x) is\ non\ linear, we\ will\ use\ iterative\ optimization\ as\ showed\ in\ the\ lecture (NG).$$

$$Linearization\ about\ \bar{x}:$$

$$x = \bar{x} + \Delta x$$

$$\hat{x}_1 = argmin\big(J(\bar{x} + \Delta x)\big)$$

$$x - \hat{x}_0 = \bar{x} + \Delta x - \hat{x}_0 = \Delta x + (\bar{x} - \hat{x}_0)$$

$$z - h(x) = z - h(\bar{x} - \Delta x) \approx z - h(\bar{x}) - H\Delta x$$

$$J(\bar{x} + \Delta x) = \left|\left|\Delta x + (\bar{x} - \hat{x}_0)\right|\right|_{\Sigma_0}^2 + \left|\left|H\Delta x + h(\bar{x}) - z\right|\right|_{\Sigma_v}^2 =$$

$$= \left|\left|\Sigma_0^{-\frac{1}{2}}\big(\Delta x + (\bar{x} - \hat{x}_0)\big)\right|\right|^2 + \left|\left|\Sigma_v^{-\frac{1}{2}}(H\Delta x + h(\bar{x}) - z)\right|\right|^2$$

$$A = \begin{pmatrix} \Sigma_0^{-\frac{1}{2}} \\ \Sigma_v^{-\frac{1}{2}} H \end{pmatrix}, b = \begin{pmatrix} \Sigma_0^{-\frac{1}{2}}(\bar{x} - \hat{x}_0) \\ \Sigma_v^{-\frac{1}{2}}(h(\bar{x}) - z) \end{pmatrix}$$

$$\hat{x}_1 = argmin\left|\left|A\Delta x + b\right|\right|^2$$

$$\boxed{\begin{matrix} \hat{x}_1 = (A^T A)^{-1} A^T b \\ \Sigma_1 = A^T A \\ p(x|z_1) = N(\hat{x}_1, \Sigma_1) \end{matrix}}$$

(d) A second measurement, z_2, is obtained. Assuming $p\,(x|z_1) = N\,(\hat{x}_1, \Sigma_1)$ from last clause is given, derive expressions for $p\,(x|z_1, z_2) = N\,(\hat{x}_2, \Sigma_2)$.

$$p(x|z_1, z_2) \underset{BR}{=} \frac{p(x|z_1)p(z_2|x, z_1)}{p(z_2|z_1)} = \frac{p(x|z_1)p(z_2|x)}{p(z_2)} = \eta p(x|z_1)p(z_2|x)$$

$$J(x) = \left|\left|x - \hat{x}_1\right|\right|_{\Sigma_1}^2 + \left|\left|z_2 - h(x)\right|\right|_{\Sigma_v}^2$$

$$\hat{x}_2 = argmax\big(p(x|z_1, z_2)\big) = argmin\Big(-\log\big(\eta p(x|z_1)p(z_2|x)\big)\Big) = \mathrm{argmin}(J(x))$$

$$since\ h(x)\ is\ non\ linear, we\ will\ use\ iterative\ optimization\ as\ showed\ in\ the\ lecture.$$
$$Linearization\ about\ \bar{x}:$$

$$x = \bar{x} + \Delta x$$

$$\hat{x}_2 = argmin\big(J(\bar{x} + \Delta x)\big)$$

$$x - \hat{x}_1 = \bar{x} + \Delta x - \hat{x}_1 = \Delta x + (\bar{x} - \hat{x}_1)$$

$$z_2 - h(x) = z_2 - h(\bar{x} - \Delta x) \approx z_2 - h(\bar{x}) - H\Delta x$$

$$J(\bar{x} + \Delta x) = \left|\left|\Delta x + (\bar{x} - \hat{x}_1)\right|\right|_{\Sigma_1}^2 + \left|\left|H\Delta x + h(\bar{x}) - z_2\right|\right|_{\Sigma_v}^2$$

$$\left|\left|\Sigma_1^{-\frac{1}{2}}\big(\Delta x + (\bar{x} - \hat{x}_1)\big)\right|\right|^2 + \left|\left|\Sigma_v^{-\frac{1}{2}}(H\Delta x + h(\bar{x}) z_2)\right|\right|^2$$

$$A = \begin{pmatrix} \Sigma_1^{-\frac{1}{2}} \\ \Sigma_v^{-\frac{1}{2}} H \end{pmatrix}, b = \begin{pmatrix} \Sigma_1^{-\frac{1}{2}}(\bar{x} - \hat{x}_1) \\ \Sigma_v^{-\frac{1}{2}}(h(\bar{x}) - z_2) \end{pmatrix}$$

$$\hat{x}_2 = argmin\left|\left|A\Delta x + b\right|\right|^2$$

$$\boxed{\begin{matrix} \hat{x}_2 = (A^T A)^{-1} A^T b \\ \Sigma_2 = A^T A \\ p(x|z_1, z_2) = N(\hat{x}_2, \Sigma_2) \end{matrix}}$$

Question 3 : Consider a multivariate random variable $x_k \in \mathbb{R}^n$. with the following state transition mode and a standard observation model as in exercise 2.

$$x_{k+1} = f(x_k, u_k) + w_k \ , w_k \sim N(0, \Sigma_w)$$

(a) Write an expression for the motion mode $p(x_k | x_{k-1}, u_{k-1})$.

$$p(x_k|x_{k-1}, u_{k-1}) =$$
$$= \frac{1}{\sqrt{\det(2\pi\Sigma_w)}} \exp\left(-\frac{1}{2}\left(x_k - f(x_{k-1}, u_{k-1})\right)^T \Sigma_w \left(x_k - f(x_{k-1}, u_{k-1})\right)\right) =$$
$$= \boxed{\frac{1}{\sqrt{\det(2\pi\Sigma_w)}} \exp\left(-\frac{1}{2}||x_k - f(x_{k-1}, u_{k-1})||^2_{\Sigma_w}\right)}$$

(b) Assume the robot executes action u_0 and then acquires a measurement z_1. Write an expression for the a posteriori pdf $p(x_1|z_1, u_0)$ in terms of the prior, motion and observation models.

$$The\ prior - p(x)$$

$$The\ motion\ modle - p(x_k|x_{k-1}, u_{k-1})$$

$$The\ observation\ modle - p(z_1|x_1)$$

$$p(x_1|z_1, u_0) \underset{BR}{=} \frac{p(x_1|u_0)p(z_1|x_1, u_0)}{p(z_1|u_0)} \underset{Indep.}{=} \frac{p(x_1|u_0)p(z_1|x_1)}{p(z_1)} = \eta p(x_1|u_0)p(z_1|x_1)$$

$$p(x_1|u_0) \underset{MR}{=} \int p(x_1, x_0|u_0)dx_0 \underset{CR}{=} \int p(x_1|x_0, u_0)p(x_0|u_0)dx_0 \underset{indep.}{=} \int p(x_1|x_0, u_0)p(x_0)dx_0$$

$$\boxed{p(x_1|z_1, u_0) = \eta p(z_1|x_1) \int p(x_1|x_0, u_0)p(x_0)dx_0}$$

(c) In the same setting, consider the a posteriori pdf over the joint state

$$x = x_{0:1} = \binom{x_0}{x_1}, with\ x_i \in \mathbb{R}^n$$

Show that calculating the maximum a posteriori (MAP) estimate for x is equivalent to solving a non-linear least squares problem.

$$In\ this\ case\ we\ use\ smoothing\ instead\ of\ marginlization.$$

$$As\ showed\ in\ the\ lecture : p(x_{0:1}|u_{x:k}, z_{1:k}) = \eta p(x_0) \prod_{i=1}^{k} p(x_i|x_{i-1}, u_{i-1})p(z_i|x_i)$$

$$p(x_{0:1}|u_0, z_1) = \eta p(x_0)p(x_1|x_0, u_0)p(z_1|x_1)$$

$$We\ will\ find\ the\ MAP\ as\ follows$$

$$x^* = argmax\left(p(x_{0:1}|u_0, z_1)\right) = argmax\left(\eta p(x_0)p(x_1|x_0, u_0)p(z_1|x_1)\right) =$$
$$= argmin\left(-\log\left(p(x_0)p(x_1|x_0, u_0)p(z_1|x_1)\right)\right) = argmin(J(x))$$

$$where\ J(x) = ||x_0 - \hat{x}_0||^2_{\Sigma_0} + ||z_1 - h(x_1)||^2_{\Sigma_v} + ||x - f(x_{k-1}, u_{k-1})||^2_{\Sigma_w}$$

$$Like\ in\ Q2\ we\ will\ use\ linearization$$

$$x_k = x_{k-1} + \Delta x$$

$$x^* = argmin\big(J(x_{k-1} + \Delta x)\big)$$

$$x_k - \hat{x}_0 = x_{k-1} + \Delta x - \hat{x}_0 = \Delta x + (x_{k-1} - \hat{x}_0)$$

$$z_1 - h(x_k) = z_1 - h(x_{k-1} - \Delta x) \approx z_1 - h(x_{k-1}) - H\Delta x$$

$$J(x_{k-1} + \Delta x) = ||\Delta x + (x_{k-1} - \hat{x}_0)||^2_{\Sigma_0} + ||H\Delta x + h(x_{k-1}) - z_1||^2_{\Sigma_v} + \left|\left|\Delta x + \big(x_{k-1} - f(x_{k-1}, u_{k-1})\big)\right|\right|^2_{\Sigma_w}$$

$$\left|\left|\Sigma_0^{-\frac{1}{2}}(\Delta x + (x_{k-1} - \hat{x}_0))\right|\right|^2 + \left|\left|\Sigma_v^{-\frac{1}{2}}(H\Delta x + h(x_{k-1})z_1)\right|\right|^2 + \left|\left|\Sigma_w^{-\frac{1}{2}}\Big(\Delta x + \big(x_{k-1} - f(x_{k-1}, u_{k-1})\big)\Big)\right|\right|^2$$

$$A = \begin{pmatrix} \Sigma_0^{-\frac{1}{2}} \\ \Sigma_v^{-\frac{1}{2}}H \\ \Sigma_w^{-\frac{1}{2}} \end{pmatrix}, b = \begin{pmatrix} \Sigma_0^{-\frac{1}{2}}(x_{k-1} - \hat{x}_0) \\ \Sigma_v^{-\frac{1}{2}}(h(x_{k-1}) - z_1) \\ \Sigma_w^{-\frac{1}{2}}\big(x_{k-1} - f(x_{k-1}, u_{k-1})\big) \end{pmatrix}$$

$$\boxed{x^* = argmin\left|\left|A\Delta x + b\right|\right|^2}$$

(d) Assume the a posteriori pdf over the joint state$x_{0:1}$ is given in covariance and information forms as

$$p(x_{0:1}|u_0, z_1) = N(\hat{x}_{0:1}, \Sigma_{0:1}) = N^{-1}(\hat{\eta}_{0:1}, I_{0:1})$$
$$\Sigma_{0:1} = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} \\ \Sigma_{01}^T & \Sigma_{11} \end{bmatrix}, I_{1:0} = \begin{bmatrix} \mathrm{I}_{00} & \mathrm{I}_{01} \\ \mathrm{I}_{01}^T & \mathrm{I}_{11} \end{bmatrix}$$

Indicate the dimensionality of the covariance matrix $\Sigma_{0:1}$ and of its components. We are interested in the marginal pdf over the state x_1, while marginalizing out the past state x_0. Write expressions for the marginal covariance and information matrices, Σ_{01} and I_{01} , over the state x_1 such that:

$$p(x_{0:1}|u_0, z_1) = N(\times, \Sigma_1') = N^{-1}(\times, I_1')$$

$$x_k \in \mathbb{R}^n \rightarrow x_{0:1} \in \mathbb{R}^{2n} \rightarrow \Sigma_{0:1} \in \mathbb{R}^{2n\times 2n}$$

$$p(x_1|u_0, z_1) = N(\times, \Sigma_{11}) = N^{-1}(\times, I_{11} - I_{01}^T I_{00} I_{10})$$