Section 1: State Space Models

1.1 System Set-up

Consider the evolution of the state sequence $\{\mathbf{x}_k, k \in \mathbb{N}\}$ of a target given by:

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \tag{1.1}$$

where $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_x}$ is a possibly nonlinear function of the state $\mathbf{x}_{k-1}$, $\{v_{k-1}, k \in \mathbb{N}\}$ is an iid process noise sequence, n_x, n_v are dimensions of the state and process noise vectors. The objective is to recursively estimate $\mathbf{x}_k$ from measurements

$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{n}_k) \tag{1.2}$$

where $h_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_n} \to \mathbb{R}^{n_z}$ is a possibly nonlinear function, $\{n_k, k \in \mathbb{N}\}$ is an iid measurement noise sequence, n_z, n_n are dimensions of the measurement and measurement noise vectors. In particular, we seek filtered estimates of $\mathbf{x}_k$ based on the set of all available measurements $\mathbf{z}_{1:k} = \{\mathbf{z}_i, i = 1, ..., k\}$ up to time k.

Suppose the required pdf $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ at time $k-1$ is available. the prediction stage involves using the system model (1.1) to obtain the prior pdf of the state at time k via the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} \tag{1.3}$$

Note that (1.3) uses the fact that $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$ as (1.1) describes a Markov process of order one. At time step k, a measurement $\mathbf{z}_k$ becomes available, and this may be used to update the prior via the Bayes' rule:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{\int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})d\mathbf{x}_k} \tag{1.4}$$

Note the normalizing constant of (1.4) depends on the likelihood function $p(\mathbf{z}_k|\mathbf{x}_k)$ defined by the measurement model (1.2) and the known statistics of $\mathbf{n}_k$. In the update stage (1.4), the measurement $\mathbf{z}_k$ is used to modify the prior density to obtain the required posterior density of the current state. The recurrence relations (1.3) and (1.4) form the basis for the optimal Bayesian solution. The optimal Bayesian solution solves the problem of recursively calculating the exact posterior density. An optimal algorithm is a method for deducing this solution.

We can also obtain $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ recursively from $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-2})$. First, apply the Chapman-Kolmogorov equation from (1.3) and then apply Bayes' rule (1.4) to $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} = \frac{\int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{z}_{k-1}|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-2})d\mathbf{x}_{k-1}}{\int p(\mathbf{z}_{k-1}|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-2})d\mathbf{x}_{k-1}} \tag{1.5}$$

1.2 Copulas in 2-Dimensions

1.2.1 Definition

Let $I^2 = [0,1] \times [0,1]$ be the unit square. $B = [u_1, u_2] \times [v_1, v_2]$, with $u_1 \le u_2, v_1 \le v_2$ be a rectangular region within I^2. Let $A(u,v)$ be a function from I^2 to I. Lets define A-volume of the region B as the value

$$V_A(B) = A(u_2, v_2) - A(u_1, v_2) - A(u_2, v_1) + A(u_1, v_1) \tag{1.6}$$

Definition 1.1 Function A(u,v) is called ***quasi-monotone*** if for any rectangular area B in I^2 its A-volume is nonnegative.

Definition 1.2 Function A(u,v) is called ***grounded*** in I^2 its $A(0,v) = A(u,0) = 0 \; \forall \; u, v \in I$.

Lemma 1.3 Any grounded nonnegative quasi-monotone function on I^2 is increasing in each argument.

Definition 1.4 Function $C : I^2 \to I$ is called a ***copula*** if it is *grounded*, *quasi-monotone*, and $\forall \; u, v \in [0,1], C(u,1) = u, C(1,v) = v$.

Definition 1.5 Let $c(u,v) = \frac{\partial^2 C}{\partial u \partial v} = \frac{\partial^2 C}{\partial v \partial u}$ exist and be continuous on I^2. Then $c(u,v)$ is a ***copula density***.

Theorem 1.6 (Sklar's Theorem) Let H be a join distribution function with marginal F and G. Then there exists a copula C s.t. $\forall \, x, y, H(x,y) = C(F(x), G(y))$. If F and G are continuous, then C is unique.

Definition 1.7 Let $C(u,v)$ be a copula defining a joint distribution with marginals u and v. The following function

$$\bar{C}(u,v) = u + v - 1 + C(1-u, 1-v) \tag{1.7}$$

is called a ***survival copula***. A survival copula is a copula because it satisfies Definition 1.4:

- $\bar{C}(u,0) = v - 1 + C(1, 1-v) = v - 1 + 1 - v = 0$
- $\bar{C}(u,1) = u + 1 - 1 + C(1-u, 0) = u$
- $V_{\bar{C}}(B) = \bar{C}(u_2, v_2) - \bar{C}(u_1, v_2) - \bar{C}(u_2, v_1) + \bar{C}(u_1, v_1) = C(1-u_2, 1-v_2) - C(1-u_2, 1-v_1) - C(1-u_1, 1-v_2) + C(1-u_1, 1-v_1) \ge 0$

Use the identity $\mathbb{P}(X \ge x) + \mathbb{P}(Y \ge y) + \mathbb{P}(X \le x, Y \le y) - \mathbb{P}(X \ge x, Y \ge y) = 1$, we can obtain

$$\mathbb{P}(X \ge x, Y \ge y) \equiv S(x,y) = \bar{C}(S_1(x), S_2(y))$$

where $S_1(x) \equiv \mathbb{P}(X > x) \equiv 1 - F_X(x)$ and $S_2(y) \equiv \mathbb{P}(Y > y) \equiv 1 - G_Y(y)$. This representation of the joint survival function as a survival copula based on marginal survival functions provides a useful alternative to the regular copula in cases when survival functions are easier or more natural to use than CDFs (e.g. exponential or Weibull distributions).

1.2.2 Simple Pair Copulas

1.2.2.1 Maximum Copula

Consider $M(u,v) = min(u,v)$ on I^2 as the ***maximum copula***. Let's check how it meets the definition of a copula. Clearly, $M(u,0) = M(0,v) = 0, M(u,1) = u, M(1,v) = v$. To check the quasi-monotone property, it is necessary to consider different cases of the location of rectangle B with respect to the diagonal $u = v$ of the unit square. There are 3 possible cases to consider:

- All 4 vertices lie on the same side of the diagonal. $V_M(B) = M(u_2, v_2) - M(u_1, v_2) - M(u_2, v_1) + M(u_1, v_1) = u_2 - u_1 - u_2 + u_1 = 0$ if B lies on the left side of the diagonal. $V_M(B) = v_2 - v_2 - v_1 + v_1 = 0$ if B lies on the right side of the diagonal.

- 3 vertices lie on the same side of the diagonal. $V_M(B) = u_2 - u_1 - u_2 + v_1 = v_1 - u_1 \geq 0$, as $(u1, v1)$ is on the right of the diagonal. Alternatively, $V_M(B) = v_2 - v_2 - v_1 + u_1 = u_1 - v_1 \geq 0$ when $(u1, v1)$ is on the left of the diagonal.

- 2 vertices lie on the same side of the diagonal. When $(u_1, v_2), (u_1, v_1)$ are on the left side,$V_M(B) = v_2 - u_1 - v_1 + u_1 = v_2 - v_1 \geq 0$. When $(u_1, v_2), (u_2, v_2)$ are on the left side, $V_M(B) = u_2 - u_1 - v_1 + v_1 = u_2 - u_1 \geq 0$.

Let X be a continuous random variable. Let $Y = \phi(X)$ be an **increasing** differentiable function of X. According to Sklar's theorem, there exists a copula which models the dependence between X and Y. It turns out this copula is the Maximum Copula. To see this, let H be the joint CDF with marginal F and G. $H(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x, \phi(X) \leq y) = \mathbb{P}(X \leq x, X \leq \phi^{-1}(y)) = min(F_X(x), F_X(\phi^{-1}(y)) = min(F_X(x), G_Y(y)) = C(F_X(x), G_Y(y))$. This is analogous to perfect correlation when both X and Y are linear, but copula is able to capture dependence when X and Y are none-linear as well.

1.2.2.2 Minimum Copula

Consider the function $W(u, v) = max(u + v - 1, 0)$ on I^2 and call it the ***minimum copula***. Let's first check if it satisfies the definition of a copula.

- $W(u, 0) = W(0, v) = 0$.

- $W(u, 1) = u, W(1, v) = v$.

- Need to verify that $V_W(B) = W(u_2, v_2) - W(u_1, v_2) - W(u_2, v_1) + W(u_1, v_1) \geq 0$. First, note if $W(u_1, v_2) > 0$ or $W(u_2, v_1) > 0$, then $W(u_2, v_2) > 0$. Let's assume the worst case that both $W(u_1, v_2) > 0$ and $W(u_2, v_1) > 0$, then $V_W(B) = (u_2 + v_2 - 1) - (u_2 + v_1 - 1) - (u_1 + v_2 - 1) + W(u_1, v_1) = 1 - u_1 - v_1 + max(0, u_1 + v_1 - 1) \geq 0$.

Analogous to Maximum Copula, let $Y = \phi(X)$ be an **decreasing** differentiable function of X. According to Sklar's theorem, there exists a copula which models the dependence between X and Y. It turns out this copula is the Minimum Copula. To see this, let H be the joint CDF with marginal F and G. $H(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x, \phi(X) \leq y) = \mathbb{P}(X \leq x, X \geq \phi^{-1}(y)) = max(0, \mathbb{P}(X \leq x) + \mathbb{P}(X \geq \phi^{-1}(y)) - 1) = max(0, F_X(x) + G_Y(y) - 1) = C(F_X(x), G_Y(y))$. Thus, the minimum copula provides another model for direct negative relationship between X and Y, which could be both linear and nonlinear.

Theorem 1.8 (Frechet-Hoeffding bounds) For any copula C and any $u, v \in [0, 1]$, the following inequality holds:

$$W(u, v) \leq C(u, v) \leq M(u, v) \tag{1.8}$$

Proof: For a copula $C(F_X(x), G_Y(y))$ be able to represent joint CDF $P(X \leq x, Y \leq y)$, $C(F_X(x), G_Y(y))$ must be less than or equal to $min(F_X(x), G_Y(y))$, or $M(F_X(x), G_Y(y))$.
Now consider a rectangle B with vertices $(1, 1), (u, 1), (1, v), (u, v)$. For this rectangle and for any copula $C(u, v)$ it is true that $V_C(B) = 1 - u - v + C(u, v) \geq 0$, so $C(u, v) \geq u + v - 1$ and $C(u, v) \geq 0$ always hold, implying $C(u, v) \geq W(u, v)$. ■

1.2.2.3 Product Copula

Consider the function $P(u,v) = uv$ on I^2 and call it the ***product copula***. Let's first check if it satisfies the definition of a copula.

- $W(u,0) = W(0,v) = 0$.
- $W(u,1) = u, W(1,v) = v$.
- $V_P(B) = P(u_2,v_2) - P(u_1,v_2) - P(u_2,v_1) + P(u_1,v_1) = u_2v_2 - u_1v_2 - u_2v_1 + u_1v_1 = (u_2-u_1)(v_2-v_1) \geq 0$

The case of two independent random variables X and Y has the product copula. To see this, again let H be the joint CDF with marginal F and G. $H(x,y) = \mathbb{P}(X \leq x, Y \leq y) = F_X(x)G_Y(y) = C(F_X(x), G_Y(y))$.

1.2.2.4 FGM Copula

The parametric family of copulas known as ***Farlie-Gumbel-Morgenstern (FGM)*** class has the form:

$$C(u,v) = uv(1 + \alpha(1-u)(1-v)),\ \alpha \in [0,1] \tag{1.9}$$

This is an easy and practical choice for modeling weaker dependence, but it is not particularly suitable for stronger dependence.

1.2.3 Elliptical Copulas

In this section we require the copulas to demonstrate elliptical symmetry, which will cause the copula functions to be symmetric with respect to the diagonal $u = 1 - v$. This will result in the identity $C(u,v) = \bar{C}(u,v)$, so that an elliptically symmetric copula coincides with its survival version. We will consider the class of bivariate elliptical distributions $Q_\rho(s,t)$ defined by their density functions

$$q_\rho(s,t) = \frac{k^2}{\sqrt{1-\rho^2}}\, g\left(\frac{s^2 - 2\rho st + t^2}{1-\rho^2}\right)$$

Here $\rho \in (-1,1)$, function $g : \mathbb{R} \to \mathbb{R}^+$ is such that $\int_{-\infty}^{\infty} g(t)dt < \infty$, and k is the normalizing constant. Consider also

$$Q(t) = \int_{-\infty}^{\infty} q_0(s,t)ds = \int_{-\infty}^{\infty} q_0(t,s)ds$$

as the marginal distribution of the first and second components of the vector (s,t) corresponding to $\rho = 0$ with symmetric density $q(t) = kg(t^2)$.

1.2.3.1 Method of Inverses

One of the possible methods of building an elliptically symmetric copula using an elliptic distribution $Q_\rho(s,t)$ is the method of inverses. Using the fact that for U and V independently uniformly distributed on $[0,1]$, the inverse transforms $Q^{-1}(U)$ and $Q^{-1}(V)$ are two independent random variables with the same CDF $Q(t)$, we will define an elliptical copula for any $u, v \in [0,1]$ as

$$C_\rho(u,v) = Q_\rho(Q^{-1}(u), Q^{-1}(v))$$

The copula density will assume the form

$$c_\rho(u,v) = \frac{q_\rho(s,t)}{q(s)q(t)}, \; s = Q^{-1}(u), t = Q^{-1}(v)$$

This construction is useful because it allows for effective separation of marginal CDF $u = F(x)$ and $v = G(y)$, which could be chosen freely. This way a bivariate joint distribution $H(x,y)$ with margins $F(x)$ and $G(y)$ can be modeled as

$$H(x,y) = Q_\rho(Q^{-1}(F(x)), Q^{-1}(G(y))) \tag{1.10}$$

where the parameter ρ is responsible for the strength of dependence.

1.2.3.2 Gaussian Copula

We obtain the ***Gaussian copula*** if $H(x,y)$ in equation (1.10) is constructed with bivariate normal CDF $\Phi_\rho(s,t)$ with zero means, unit variances, and correlation ρ. Let $\Phi(t)$ be the standard normal CDF. We have:

$$C_\rho(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{1.11}$$

Denoting $s = \Phi^{-1}(u)$, $t = \Phi^{-1}(v)$, thus $\Phi(s) = u$, $\phi(s)ds = du$, $\frac{ds}{du} = \frac{1}{\phi(s)}$. The desnity of the Gaussian copula can be written as:

$$c_\rho(u,v) = \frac{\partial^2 C_\rho(u,v)}{\partial u \partial v} = \frac{\partial^2 \Phi_\rho(s,t)}{\partial s \partial t} \cdot \frac{\partial s}{\partial u} \cdot \frac{\partial t}{\partial v} = \frac{\phi_\rho(s,t)}{\phi(s)\phi(t)} \tag{1.12}$$

where $\phi_\rho(s,t)$ is the density function for $\Phi_\rho(s,t)$ and $\phi(t)$ is the density function for $\Phi(t)$. The explicit expression for (1.12) is:

$$c_\rho(s,t) = \frac{1}{\sqrt{1-\rho^2}}\, exp\left(-\frac{1}{2}\frac{\rho^2 s^2 - 2\rho st + \rho^2 t^2}{1-\rho^2}\right) \tag{1.13}$$

For modeling joint distributions we can combine Gaussian copula with any marginal distributions $u = F(x)$ and $v = G(y)$. The idea of Gaussian copula is to transform two random variables X and Y with respective CDFs F and G into standard normal variables $S = \Phi^{-1}(F(X))$ and $T = \Phi^{-1}(G(Y))$. Then the dependence between X and Y is expressed in terms of the dependence structure of their normal transformations S and T, reducing the dependence to linear correlation.

1.2.3.3 t-Copula

In applications it is often necessary to model heavy-tailed multivariate distributions and tails of the joint distributions. In this situation, as in the one-dimensional case, the Student t-copula may be used instead of the Gaussian copula. If we use bivariate Student t-distribution with η degrees of freedom and correlation coefficient ρ in equation (1.10), we obtain the ***t-copula***. As in the Gaussian copula, the choice of an elliptical copula model does not prescribe the choice of marginals. They might be chosen separately.

$$C_{\eta\rho}(u,v) = T_{\eta\rho}(T_\eta^{-1}(u), T_\eta^{-1}(v)) \tag{1.14}$$

For inverse transform we use two univariate t-distribution with η degrees of freedom. Let us denote $s = T_\eta^{-1}(u)$, $t = T_\eta^{-1}(v)$. The density function would be in the form:

$$c_{\eta\rho}(u,v) = \frac{\partial^2 C_{\eta\rho}(u,v)}{\partial u \partial v} = \frac{\partial^2 T_{\eta\rho}(s,t)}{\partial s \partial t} \cdot \frac{\partial s}{\partial u} \cdot \frac{\partial t}{\partial v} = \frac{\psi_{\eta\rho}(s,t)}{\psi_\eta(s)\psi_\eta(t)} \tag{1.15}$$

where $\psi_{\eta\rho}(s,t)$ is the bivariate t-distribution density and $\psi_\eta(s)$ is the univariate t-distribution density with the same degrees of freedom. Explicitly, the copula density can be expressed as:

$$c_{\eta\rho}(u,v) = \frac{\Gamma(\frac{\eta+2}{2})\Gamma(\frac{\eta}{2})}{\sqrt{1-\rho^2}\Gamma^2(\frac{\eta+1}{2})} \cdot \frac{\left((1+\frac{s^2}{\eta})(1+\frac{t^2}{\eta})\right)^{\frac{\eta+1}{2}}}{\left(1+\frac{s^2+t^2-2\rho st}{\eta(1-\rho^2)}\right)^{\frac{\eta+2}{2}}} \tag{1.16}$$

For modeling joint distributions we can combine t-copula with any marginal distributions $u = F(x)$ and $v = G(y)$.

1.3 Copulas in State Space Model Estimations

From Equation (1.5), we have the form

$$p_n(x) = \frac{\int f(x|y)p(z|y)p_{n-1}(y)dy}{\int p(z|y)p_{n-1}(y)dy} = \frac{p(x,z)}{p(z)}$$

multiply $f(x)$ to both top and bottom, and use the fact that $c(F_A(a), F_B(b)) = \frac{f_{AB}(a,b)}{f_A(a)f_B(b)}$, we can then use copula in the estimation of $p_n(x)$:

$$p_n(x) = \frac{p(x,z)f(x)}{p(z)f(x)} = c(F_X(x), F_Z(z))f(x) \tag{1.17}$$

where $f(x)$ can be obtained through $\int f(x|y)p_{n-1}(y)dy$.

Using Equation (1.17), we can obtain the CDF $P_n(x)$:

$$P_n(x) = \int_{-\infty}^{x} p_n(w)dw = \int_{-\infty}^{x} c(F_X(w), F_Z(z))f(w)dw = \frac{\partial C(F_X(x), F_Z(z))}{\partial F_Z(z)} \tag{1.18}$$

1.3.1 Gaussian Copula Example

Initialize $x_0 \sim N(0,1)$. Let $f(x_k|x_{k-1})$ be $(x_{k-1}+1)+u$, $u \sim N(0,1)$. Thus, $f(x_1)$ can be obtained through:

$$\begin{aligned} f(x_1) &= \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_1-(x_0+1))^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_0^2} dx_0 \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_1-1)^2} \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(2x_0^2-2x_0(x_1-1))} dx_0 \\ &= \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{1}{2}\frac{1}{2}(x_1-1)^2} \sim N(1,\sqrt{2}) \end{aligned} \tag{1.19}$$

Let $p(z_k|x_k)$ be $x_k + v$, $v \sim N(0,2)$:

$$\begin{aligned} p(z_0) &= \int \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\frac{1}{4}(z_0-x_0)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_0^2} dx_0 \\ &= \frac{1}{\sqrt{5}\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{1}{4}-\frac{1}{4\cdot 5})z_0^2} \int \frac{\sqrt{5/4}}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{5}{4}(x_0^2-\frac{1}{5}z_0)^2} dx_0 \\ &= \frac{1}{\sqrt{2\pi}\sqrt{5}} e^{-\frac{1}{2}\frac{1}{5}z_0^2} \sim N(0,\sqrt{5}) \end{aligned} \tag{1.20}$$

For a Gaussian Copula, plug in (1.13) into Equation (1.17), we get:

$$p_n(x) = \frac{1}{\sqrt{1-\rho^2}} exp\left(-\frac{1}{2}\frac{\rho^2 s^2 - 2\rho st + \rho^2 t^2}{1-\rho^2}\right) f(x) \tag{1.21}$$

where $s = \Phi^{-1}(F_X(x))$, $t = \Phi^{-1}(F_Z(z))$.

1.3.2 Implementation for Normal Distribution

In this section, we try to estimate densities for latent variables x_k based on observations z_{k-1} in the following state space system:

$$\begin{aligned} x_k &= x_{k-1} + 1 + \sqrt{v_x}\epsilon, \qquad \epsilon \sim N(0,1) \\ z_k &= x_k + \sqrt{v_z}\epsilon', \qquad \epsilon' \sim N(0,1) \end{aligned} \tag{1.22}$$

1.3.2.1 Kalman Filter Method

Assume $x_k|z_k \sim N(\mu_k, w_k)$, and let ϵ'' be another standard normal random variable, we have:

$$\begin{aligned} x_{k+1} &= x_k + 1 + \sqrt{v_x}\epsilon = \mu_k + \sqrt{w_k}\epsilon'' + 1 + \sqrt{v_x}\epsilon \\ &\Rightarrow p(x_{k+1}|z_k) = N(\mu_k + 1, w_k + v_x) \end{aligned} \tag{1.23}$$

Based on Kalman Filter algorithm, we have:

$$\mu_t = w_t\left(\frac{z_t}{v_z} + \frac{\mu_{t-1}+1}{v_x + w_{t-1}}\right) \qquad w_t = \frac{v_z(v_x + w_{t-1})}{v_z + v_x + w_{t-1}} \tag{1.24}$$

1.3.2.2 Grid Method

Conduct numerical integration on grids.

1. Draw samples from the state space system in Equation (1.22).
2. Lay out a grid of x, compute $f(x_1)$, $p(z_0)$, $p(x_1|z_0)$ as given in Equations (1.19), (1.20), (1.21).
3. The next marginal density $f(x_{k+1})$ can be obtained through numerical integration over a 2-dimensional grid of $[x_{k+1}, x_k]$:
$$f(x_{k+1}|z_{k-1}) = \int f(x_{k+1}|x_k)p(x_k|z_{k-1})dx_k \tag{1.25}$$
4. The marginal distribution $P(z_k)$ can be obtained through numerical integration over 1-dimensional grid of x_k:
$$P(z_k|z_{k-1}) = \int P(z_k|x_k)p(x_k|z_{k-1})dx_k \tag{1.26}$$
5. Again use Equation (1.21) to compute $p_{k+1}(x) = p(x_{k+1}|z_k)$.
6. Repeat steps 3 to 5.

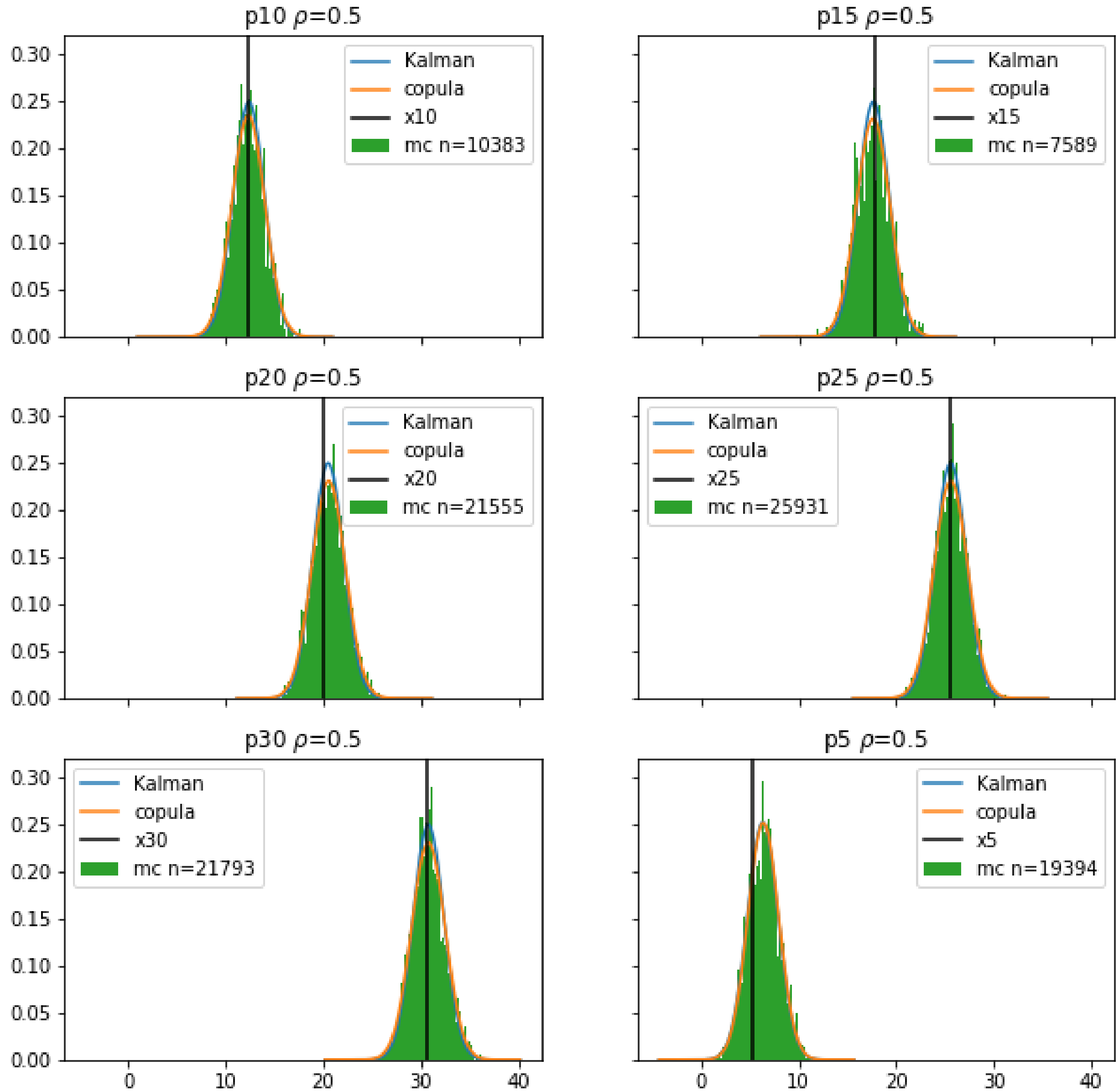
1.3.2.3 Monte Carlo Method

Simulate through Monte Carlo. Integration is evaluated with accept-reject algorithm. Here are the steps:

1. Draw samples from the state space system in Equation (1.22).
2. Select end points E_0 and E_1. Generate m uniformly distributed random number between E_0 and E_1. Call it x_u.
3. Monte Carlo integration: Compute all $f(x_u|x_k)$, where x_k represent the particles for $p(x_k|z_{k-1})$. Average over x_k to obtain $f(x_u)$.
4. Accept those x_u with probability $\frac{f(x_u)}{max(f(x_u))}$. We can do this by generating another m uniform random variables between 0 and the maximum value of $f(x_u)$. The marginal distribution $f(x_{k+1}|z_{k-1})$ is approximated by the remaining x_u.
5. Monte Carlo integration: Obtain the CDF of the accepted x_u, or $F(x_{k+1}|z_{k-1})$, by computing $F(x_u|x_k)$ and then averaging over x_k.
6. Monte Carlo integration: Obtain $P(z_k|z_{k-1})$ by averaging $P(z_k|x_k)$ over x_k.
7. Compute the copula density $c(F_X(x_{k+1}), F_Z(z_k))$, where x_{k+1} are the accepted x_u.
8. Compute $p(x_{k+1}|z_k) = c(F_X(x_{k+1}), F_Z(z_k))f(x_{k+1})$ by bootstrapping another m samples from x_u, accept them with probability $\frac{c}{max(c)}$. The accepted x_u are particles for the desired density $p(x_{k+1}|z_k)$.
9. Repeat steps 2 to 8.

1.3.2.4 Results Comparison

Because both $f(x_{k+1}|x_k)$ and $h(z_k|x_k)$ are normal, Kalman Filter represents the true density. It serves as our baseline for the efficacy of copula method in estimating hidden Markov models. Below are the density estimations for $p(x_{k+1}|z_k)$ from three different implementations: Kalman Filter, Grid Method, and Monte Carlo Method. As one can see, for the choice of $\rho = 0.5$, the copula method estimation is very close to the truth (Kalman Filter). The green bars represent density histograms of the particles resulting from the Monte Carlo method. The letter n in the graph legends denotes number of particles.


p10 ρ=0.5
Kalman
copula
x10
mc n=10383
p15 ρ=0.5
Kalman
copula
x15
mc n=7589
p20 ρ=0.5
Kalman
copula
x20
mc n=21555
p25 ρ=0.5
Kalman
copula
x25
mc n=25931
p30 ρ=0.5
Kalman
copula
x30
mc n=21793
p5 ρ=0.5
Kalman
copula
x5
mc n=19394
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
10
20
30
40

1.3.2.5 Particle Regeneration

Suppose $u = C(F_X(x_{k+1}), F_Z(z_k)))$ is given, we want to figure out how to obtain x_{k+1}. Essentially, this problem boils down to if we are given $u' = F_X(x_{k+1})$, how can we back out x_{k+1} using the equation

$$F(x_{k+1}|z_{k-1}) = \int F(x_{k+1}|x_k)p(x_k|z_{k-1})dx_k \tag{1.27}$$

With our set-up, we have a bunch of particles presenting $p(x_k|z_{k-1})$. The following outlines an algorithm that would help us obtain new particles of x_{k+1}.

1. First select several values of x_{k+1} far apart enough, compute $F(x_{k+1}|z_{k-1})$ through MC integration using Equation (1.27).
2. Use the x_{k+1} whose CDF encloses u'. Generate finer values of x_{k+1}. Further compute $F(x_{k+1}|z_{k-1})$ and save them.
3. repeat the last step until arbitrary closeness to u' is found.
4. repeat for other u's, utilize previously saved values of x_{k+1} and $F(x_{k+1}|z_{k-1})$.