תרגיל 2 | BML
מגיש: אבי כוגן | ת.ז: 205417710

שאלה 1

נתון $p(\theta|D)$ נבצע החלפת משתנים מ - α, γ לקבל הצפיפויות הנדרשות.
נשתמש בנוסחה להחלפת משתנים שראינו בתרגול 1:
נגדיר פונקציות לפי הפרמטריזציה הנתונה של y לפי כל פרמטר:

$$f : \theta \to \alpha$$
$$g : \theta \to \gamma$$

$$f(\theta) = \frac{\theta}{10} \Rightarrow f^{-1}(\alpha) = 10\alpha$$
$$g(\theta) = \theta^{1/3} \Rightarrow g^{-1}(\gamma) = \gamma^3 = (\gamma_1^3, .., \gamma_k^3)$$

$$\frac{\partial f^{-1}}{\partial \alpha} = 10$$
$$\frac{\partial g^{-1}}{\partial \gamma} = 3diag(\gamma_1^2, .., \gamma_n^2)$$

$$p_\alpha(\alpha|D) = p_\theta(f^{-1}(\alpha)|D) \cdot |\frac{\partial f^{-1}}{\partial \alpha}| = 10^k p_\theta(10\alpha|D)$$
$$p_\gamma(\gamma|D) = p_\theta(g^{-1}(\gamma)|D) \cdot |\frac{\partial g^{-1}}{\partial \gamma}| = 3^k \cdot p_\theta(\gamma^3|D) \cdot \Pi_{n=1}^k \gamma_n^2$$

שאלה 2

נבדוק עבור α:

$$
\begin{aligned}
\hat{\alpha}^{MMSE} &= E_\alpha[p_\alpha(\alpha|D)] \\
&= \int \alpha p_\alpha(\alpha|D) d\alpha \\
&\overset{Q1}{=} \int \frac{\theta}{10} p_\alpha(\alpha|D) d\alpha \\
&\overset{Q1,*}{=} \int \frac{\theta}{10} 10 p_\theta(10\alpha|D) \frac{1}{10} d\theta \\
&= \frac{1}{10} \int \theta p_\theta(\theta|D) d\theta \\
&= \frac{1}{10} \hat{\theta}^{MMSE}
\end{aligned}
$$

* עבור החלפת המשתנה באינטגרל מתקיים:

$$
\begin{aligned}
\theta &= 10\alpha \\
d\theta &= 10 d\alpha
\end{aligned}
$$

$$
y_{\hat{\alpha}}(x) \overset{def\ Q1}{=} x \cdot 10\hat{\alpha}^{MMSE} = x \cdot \hat{\theta}^{MMSE} \overset{def\ Q1}{=} y_{\hat{\theta}}(x)
$$

קיבלנו שהפתרון זהה.

נבדוק עבור γ:

$$
\begin{aligned}
\hat{\gamma}^{MMSE} &= E_\gamma[p_\gamma(\gamma|D)] \\
&= \int \gamma p_\gamma(\gamma|D) d\gamma \\
&\overset{Q1}{=} \int \gamma \cdot 3 \cdot p_\theta(\gamma^3|D) \cdot \gamma_n^2 \, d\gamma \\
&\overset{*}{=} 3 \int \gamma^3 \cdot p_\theta(\gamma^3|D) \frac{1}{3} \theta^{-2/3} d\theta \\
&\overset{\gamma^3=\theta}{=} \int \theta^{1/3} \cdot p_\theta(\theta|D) d\theta \\
&= E_\theta[\theta^{1/3}|D]
\end{aligned}
$$

* עבור החלפת המשתנה באינטגרל מתקיים:

$$\theta^{1/3} = \gamma$$
$$\frac{1}{3}\theta^{-2/3}d\theta = d\gamma$$

לכן קיבלנו:

$$y_{\hat{\gamma}}(x) \overset{def\_Q1}{=} x \cdot (\hat{\gamma}^{MMSE})^3 = x \cdot E_\theta[\theta^{1/3}|D]^3 \neq x \cdot E_\theta[\theta|D] \overset{def\_Q1}{=} y_{\hat{\theta}}(x)$$

קיבלנו שהפתרון שונה.

סה"כ קיבלנו $y_{\hat{\theta}}(x) = y_{\hat{\alpha}}(x)$,$y_{\hat{\theta}}(x) \neq y_{\hat{\gamma}}(x)$

שאלה 3

$$p(\theta, D_1, D_2) \overset{Bayes\ law}{=} p(D_1, D_2|\theta)p(\theta)$$
$$\overset{*}{=} p(D_1|\theta)p(D_2|\theta)p(\theta)$$

* מהנתון מתקיים שבהנתן θ כל התצפיות ב"ת כי $\eta \sim N(0, I\sigma^2)$, כלומר $D_1 \perp D_2$.

שאלה 4

$$p(\theta|D_1, D_2) = \frac{p(\theta, D_1, D_2)}{p(D_1, D_2)} = \frac{p(D_1|\theta)p(D_2|\theta)p(\theta)}{p(D_1, D_2)}$$
$$\overset{*}{\propto} p(D_1|\theta)p(D_2|\theta)p(\theta)$$

* כאשר מניחים ש $p(D_1, D_2)$ קבוע.
קיבלנו ש-$p(\theta|D_1, D_2)$ פרופרציאונלי למכפלת צפיפויות של גאוסיאניים, לכן זוהי גם צפיפות גאוסיאנית,
נמצא לה את התוחלת והשונות בעזרת המהנלוביס.
נסמן Y_i - וקטור המטרה, בD_i הדאטה וH_i את ה$design\ matrix$ המתאימה,

ראינו בהרצאה שמתקיים $D_i|\theta \sim N(H_i\theta, \sigma^2)$, לכן נקבל:

$$\begin{aligned}
\Delta &= \frac{1}{2}[(\theta-\mu)^T\Sigma^{-1}(\theta-\mu) + \frac{1}{\sigma^2}||Y_1 - H_1\theta||_2^2 + \frac{1}{\sigma^2}||Y_2 - H_2\theta||_2^2] \\
\frac{\partial\Delta}{\partial\theta} &= \frac{\cancel{1}}{\cancel{2}}[\cancel{2}\cdot\Sigma^{-1}(\theta-\mu) - \frac{\cancel{2}}{\sigma^2}H_1^T(Y_1 - H_1\theta) - \frac{\cancel{2}}{\sigma^2}H_2^T(Y_2 - H_2\theta)] \\
&= (\Sigma^{-1} + \frac{1}{\sigma^2}H_1^TH_1 + \frac{1}{\sigma^2}H_2^TH_2)\theta - (\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_1^TY_1 + \frac{1}{\sigma^2}H_2^TY_2)
\end{aligned}$$

וכמו שראינו בתרגול בטריק הנגזרת מתקיים:

$$\theta|D_1, D_2 \sim N((\Sigma^{-1} + \frac{1}{\sigma^2}H_1^TH_1 + \frac{1}{\sigma^2}H_2^TH_2)^{-1}(\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_1^TY_1 + \frac{1}{\sigma^2}H_2^TY_2), (\Sigma^{-1} + \frac{1}{\sigma^2}H_1^TH_1 + \frac{1}{\sigma^2}H_2^TH_2)$$

שאלה 5

ממשפט שהוכחנו בהרצאה 4 מתקיים:

$$\begin{aligned}
\mu_{\theta|D_i} &= \Sigma_{\theta|D_i}(\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_i^TY_i) \\
\Sigma_{\theta|D_i} &= (\frac{1}{\sigma^2}H_i^TH_i + \Sigma^{-1})^{-1}
\end{aligned}$$

בנוסף מתקיים שהאומד $MMSE$ ל$\theta|D_1, D_2$ הוא התוחלת של $\theta|D_1, D_2$, לכן נשכתב את התוחלת לפי הפריור על θ ו $\mu_{\theta|D_i}, \Sigma_{\theta|D_i}$:

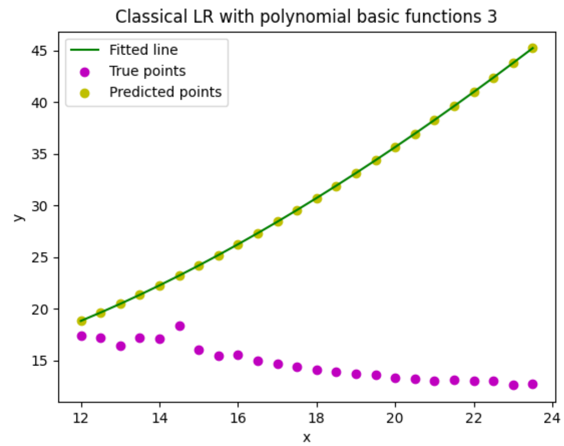
$$\begin{aligned}
\Sigma^{-1} + \frac{1}{\sigma^2}H_1^TH_1 + \frac{1}{\sigma^2}H_2^TH_2 =& \underbrace{\Sigma^{-1} + \frac{1}{\sigma^2}H_1^TH_1}_{*} + \underbrace{\Sigma^{-1} + \frac{1}{\sigma^2}H_2^TH_2}_{**} - \Sigma^{-1} \\
=& \underbrace{\Sigma_{\theta|D_1}^{-1}}_{*} + \underbrace{\Sigma_{\theta|D_2}^{-1}}_{**} - \Sigma^{-1}
\end{aligned}$$

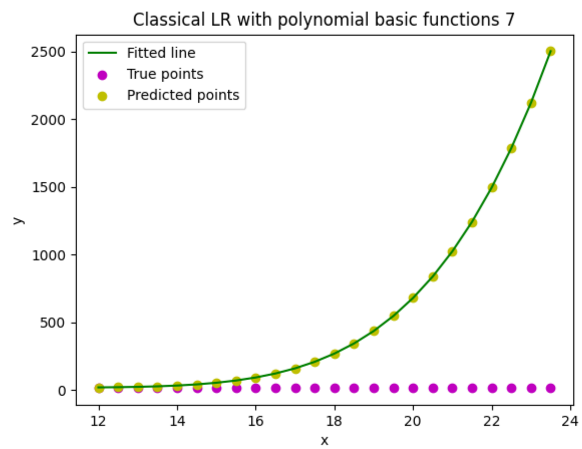
$$\begin{aligned}
\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_1^TY_1 + \frac{1}{\sigma^2}H_2^TY_2 =& \underbrace{\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_1^TY_1}_{*} + \underbrace{\Sigma^{-1}\mu + \frac{1}{\sigma^2}H_2^TY_2}_{**} - \Sigma^{-1}\mu \\
=& \underbrace{\Sigma_{\theta|D_1}^{-1}\mu_{\theta|D_1}}_{*} + \underbrace{\Sigma_{\theta|D_2}^{-1}\mu_{\theta|D_2}}_{**} - \Sigma^{-1}\mu
\end{aligned}$$

$$\Rightarrow \theta|D_1, \hat{D}_2^{MMSE} = (\Sigma_{\theta|D_1}^{-1} + \Sigma_{\theta|D_2}^{-1} - \Sigma^{-1})^{-1}(\Sigma_{\theta|D_1}^{-1}\mu_{\theta|D_1} + \Sigma_{\theta|D_2}^{-1}\mu_{\theta|D_2} - \Sigma^{-1}\mu)$$

חלק מעשי

רגרסיה קלאסית עם פונק' בסיס פולינומיאלית:

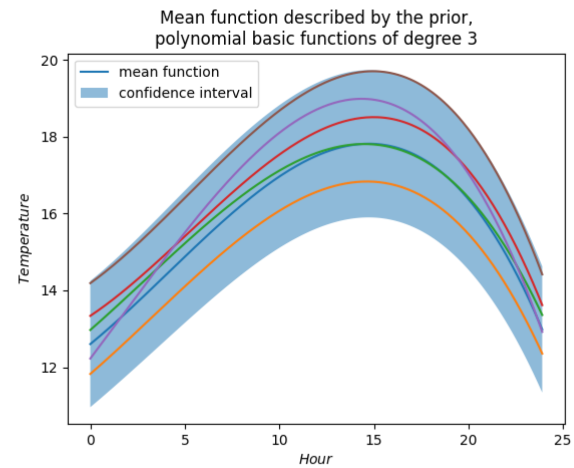


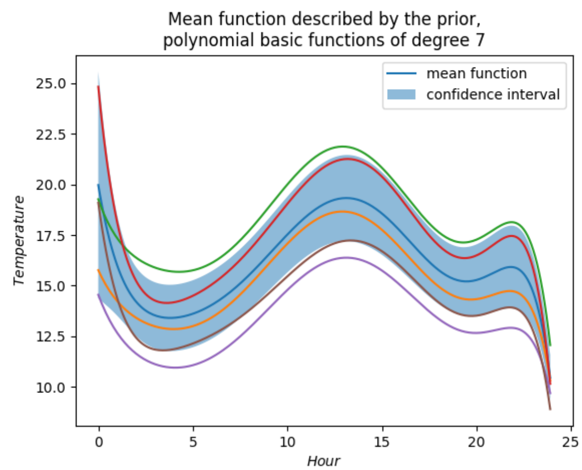



```
Average squared error with Classical LR with polynomial basic functions of degree 3 is 351.21
Average squared error with Classical LR with polynomial basic functions of degree 7 is 848707.53
```

רגרסיה בייזיאנית עם פונק' בסיס פולינומיאלית:

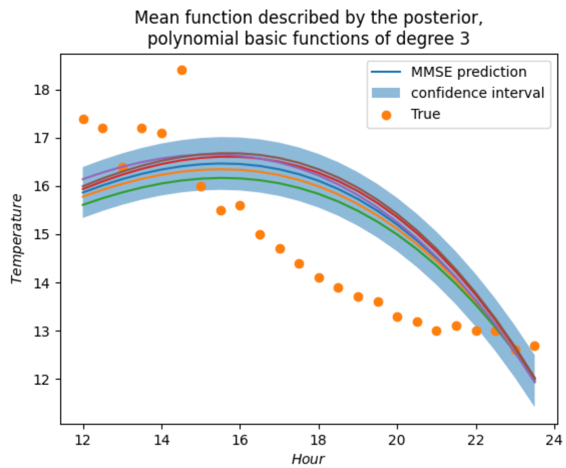
פריור:

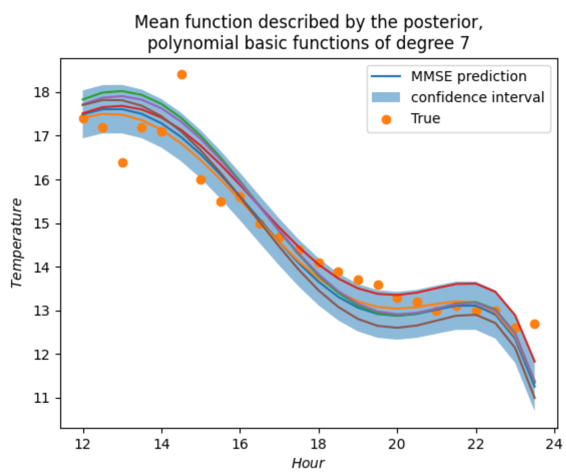




נצפה שמרבית הפונק' שיוגרלו יהיו בתוך הר"ס, ואכן ניתן לראות שהן בר"ס.

פוסטריור:

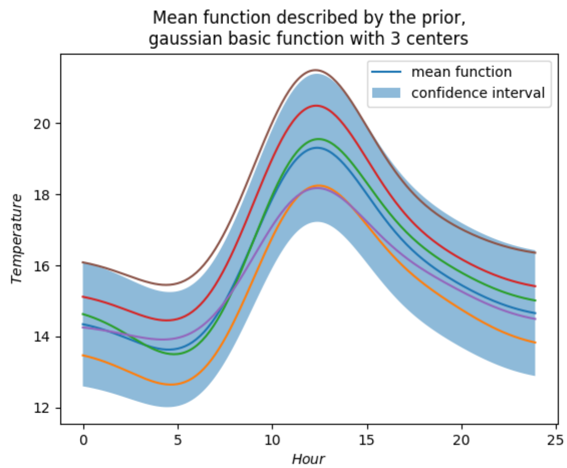


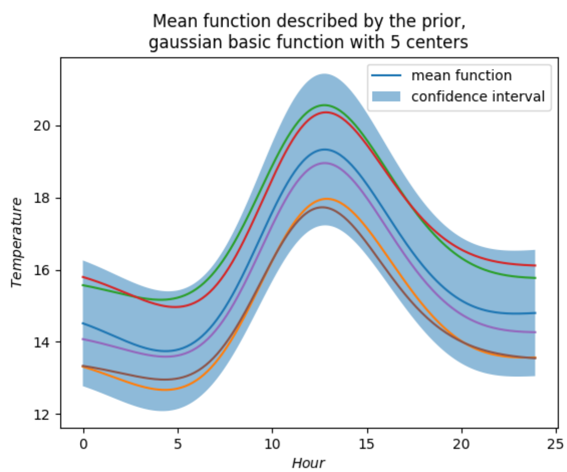


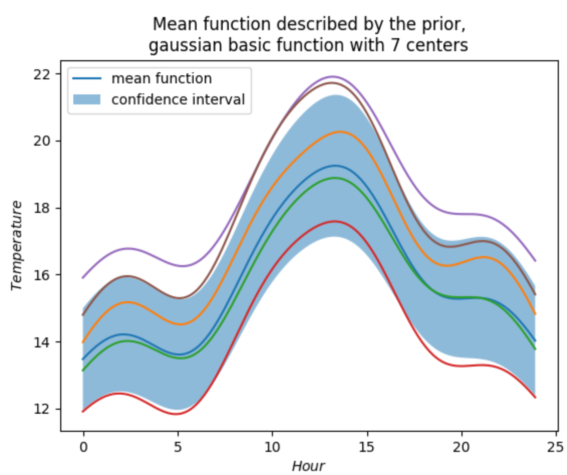

```
Average squared error with BLR with polynomial basic functions of degree 3is 1.90
Average squared error with BLR with polynomial basic functions of degree 7is 0.35
```

רגרסיה בייזיאנית עם פונק' בסיס גאוסיאניות:

פריור:

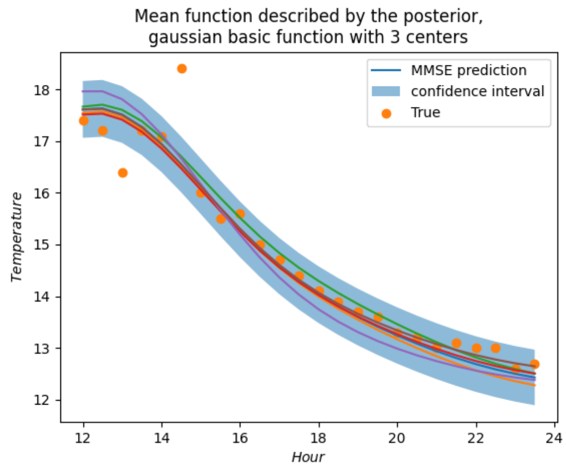

Mean function described by the prior,
gaussian basic function with 3 centers
mean function
confidence interval
Temperature
Hour

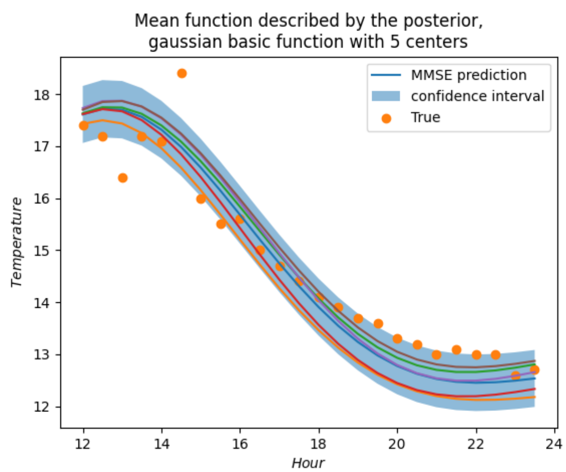

Mean function described by the prior,
gaussian basic function with 5 centers
mean function
confidence interval
Temperature
Hour

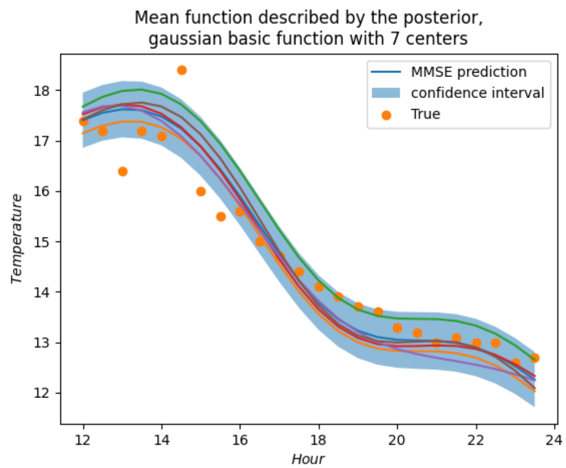

Mean function described by the prior,
gaussian basic function with 7 centers
mean function
confidence interval
Temperature
Hour

נצפה שמרבית הפונק' שיוגרלו יהיו בתוך הר"ס, ואכן ניתן לראות שעבור 3,5 מרכזים כולם בר"ס אך ב7 מרכזים 3 פונק' אינן בר"ס לכל אורכן.

פוסטריור:



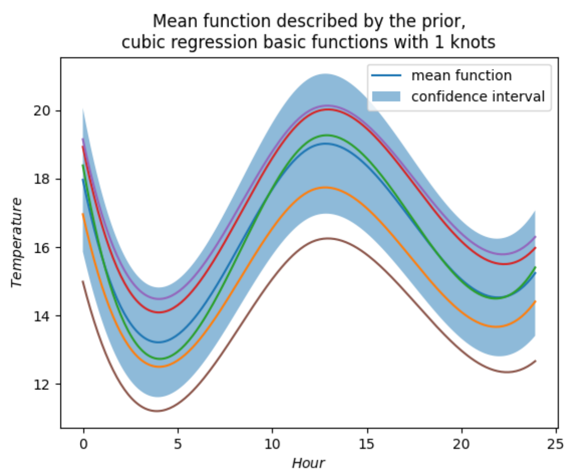


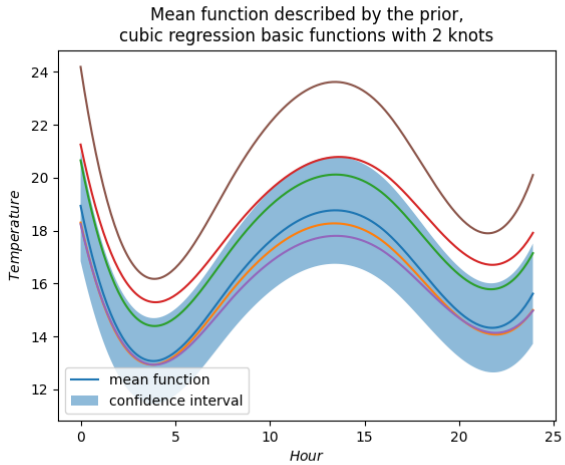


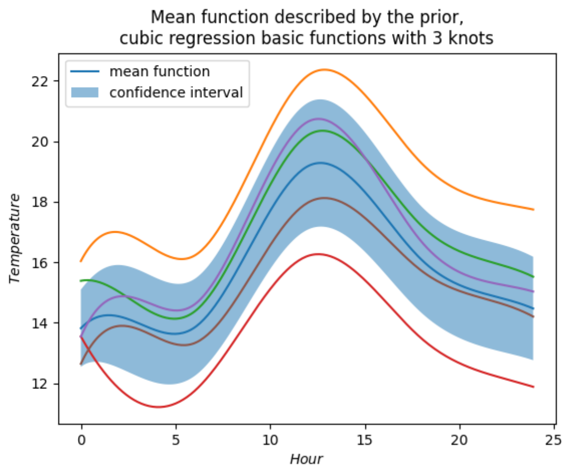

```
Average squared of BlR error with gaussian basic functions with 3 centers is 0.23
Average squared of BlR error with gaussian basic functions with 5 centers is 0.32
Average squared of BlR error with gaussian basic functions with 7 centers is 0.26
```

רגרסיה בייזיאנית עם פונק' בסיס $Cubic$:

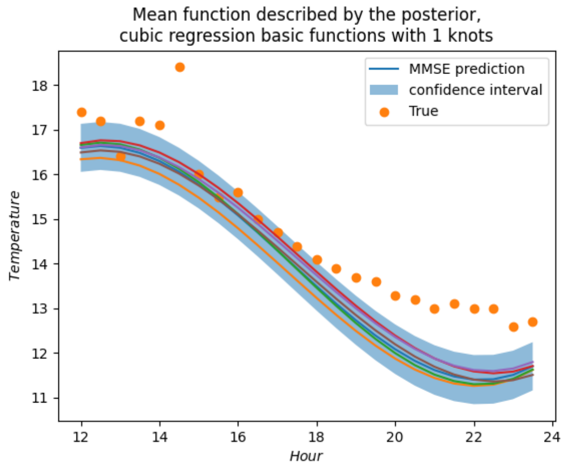
פריור:

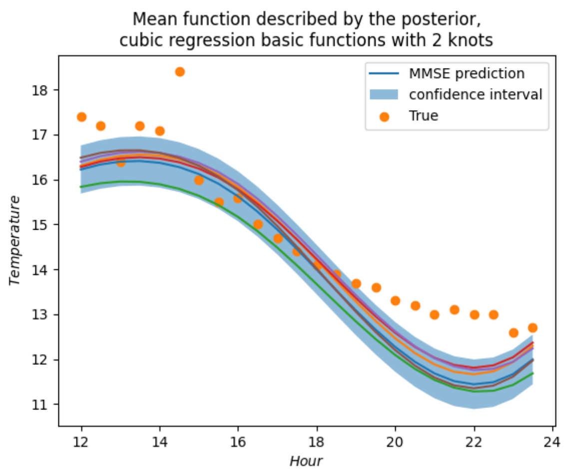


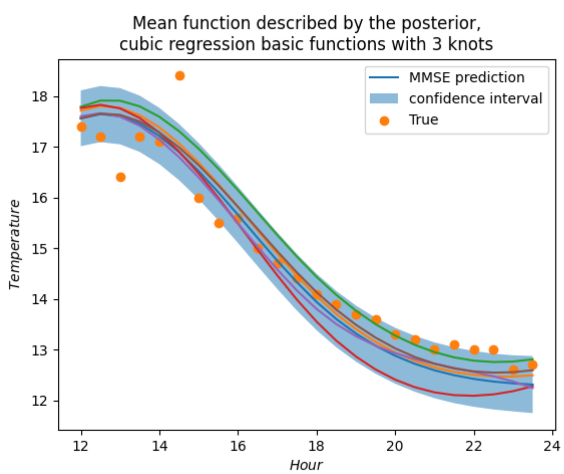




נצפה שמרבית הפונק' שיוגרלו יהיו בתוך הר"ס, ואכן ניתן לראות שעבור $knot$ 1 זה כך אך עבור 2,3 קיבלנו 2 פונק' שהן מחוץ לר"ס.


Mean function described by the posterior,
cubic regression basic functions with 1 knots
MMSE prediction
confidence interval
True
Temperature
Hour


Mean function described by the posterior,
cubic regression basic functions with 2 knots
MMSE prediction
confidence interval
True
Temperature
Hour


Mean function described by the posterior,
cubic regression basic functions with 3 knots
MMSE prediction
confidence interval
True
Temperature
Hour

```
Average squared error of BlR with cubic regression basic functions with 1 knots 1.13
Average squared error of BlR with cubic regression basic functions with 2 knots 0.94
Average squared error of BlR with cubic regression basic functions with 3 knots 0.30
```

כל השגיאות:

```
Average squared error with Classical LR with polynomial basic functions of degree 3 is 351.21
Average squared error with Classical LR with polynomial basic functions of degree 7 is 848707.53
Average squared error with BLR with polynomial basic functions of degree 3 is 1.90
Average squared error with BLR with polynomial basic functions of degree 7 is 0.35
Average squared error of BlR with gaussian basic functions with 3 centers is 0.23
Average squared error of BlR with gaussian basic functions with 5 centers is 0.32
Average squared error of BlR with gaussian basic functions with 7 centers is 0.26
Average squared error of BlR with cubic regression basic functions with 1 knots 1.13
Average squared error of BlR with cubic regression basic functions with 2 knots 0.94
Average squared error of BlR with cubic regression basic functions with 3 knots 0.30
```