תרגיל 4 | BML
מגיש: אבי כוגן | ת.ז: 205417710

תיאורטי

שאלה 1

עבור p_2:

מתקיים $\beta = \alpha \to \infty$ לכן נקבל $k(x_i,x_j) = e^{-\beta||x_i-x_j||^2} \overset{\beta\to\infty}{=} \begin{cases}1 & x_i=x_j\\ 0 & o.w\end{cases}$ לכן

$p(\mathbf{f}) \sim N(0, I_M)$

עבור p_1:

מתקיים $\alpha \to \infty \Leftarrow \beta = \frac{1}{\alpha} \to 0$ לכן נקבל $k(x_i,x_j) = e^{-\beta||x_i-x_j||^2} \overset{\beta\to 0}{=} 1$.

קיבלנו $K = \mathbf{1}\mathbf{1}^T$, לכן $p(\mathbf{f}) \sim N(0, \mathbf{1}\mathbf{1}^T)$.

שאלה 2

א.

עבור p_1:

מסעיף קודם נסיק:

$$p(\mathbf{f}|p_1) = P(\begin{pmatrix}\mathbf{f_1}\\ .\\ \mathbf{f_M}\end{pmatrix}|p_1) = \begin{cases}\frac{1}{\sqrt{2\pi}}e^{-\frac{\mathbf{f_i}^2}{2}} & \mathbf{f_i} = \mathbf{f_1}\ i\in[M]\\ o & o.w\end{cases}$$

$$\Rightarrow \begin{cases}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}} & \mathbf{f_i} = \pm 1\ \ i\in[M]\\ o & o.w\end{cases}$$

ב.

עבור p_2:

מתקיים ש-$f_i \perp f_j$ מכך ש-$p(\mathbf{f}) \sim N(0, I_M)$. ונקבל:

$$p(\mathbf{f}=v|p_2) = P(\begin{pmatrix}v_1\\ .\\ v_2\end{pmatrix}|p_2) = \frac{1}{\sqrt{2\pi}}e^{-\frac{||v||^2}{2}}$$

שאלה 3

ב2 המקרים כאשר $\alpha \to \infty$:
משאלה קודמת:

$$log(p(v|p_1)) = \begin{cases} -\frac{1}{2}(log(2\pi)) + v_1^2) & v_1 = v_i \; i \in [M] \\ 0 & o.w \end{cases}$$
$$log(p(v|p_1)) = -\frac{M}{2}(\frac{||v||^2}{M}log(2\pi))$$

עבור $\mathbf{f}_1 = (1, .., 1)^T$:

$$log(p(\mathbf{f_1}|p_1)) = -\frac{1}{2}(log(2\pi) + 1)$$
$$log(p(\mathbf{f_1}|p_2)) = M(-\frac{1}{2}(1 + log(2\pi)) = Mlog(p(\mathbf{f_1}|p_1))$$

מתקיים $log(p(\mathbf{f_1}|p_1)) < 0$ ו $M > 1$ לכן נבחר ב-p_1.

עבור $\mathbf{f}_2 = (-1, 1, .., -1, 1)^T$:

$$log(p(\mathbf{f_2}|p_1)) \to -\infty$$
$$log(p(\mathbf{f_2}|p_2)) \to Mlog(p(\mathbf{f_1}|p_1))$$

לכן נבחר בp_2.

שאלה 4

מתקיים לכל $\{x_i\}_{i=1}^N$ שמטריצת הגראם המתקבלת עבור $k(x, y) = 1$ היא $K = \mathbf{1}\mathbf{1}^T$. לכן לכל v נקבל $v^T\mathbf{1}\mathbf{1}^Tv = ||\mathbf{1}^Tv||_2^2 \geq 0$, קיבלנו ש$K$ היא PSD ולכן קרנל חוקי.

מתקיים $K = H\Sigma H^T$ עבור Σ השונות של ה-$prior$ ו-H מטריצת פונק' הבסיס מופעלות על התצפיות.

לכן נוכל להניח במקרה זה ש $\Sigma = 1$ ו-$H = \mathbf{1}$, כלומר $\theta \sim N(0, 1)$ והבעיה הפרימלית היא: $f_\theta(x) = H^T\theta$.

שאלה 5

$$\begin{aligned}
\alpha_i =& [(\mathbf{1}\mathbf{1}^T + \sigma^2 \cdot I)^{-1} y]_i \overset{*}{=} \\
=& [\sigma^{-2}(I - \cancel{\sigma^{-2}}(\frac{\cancel{\sigma^2}}{\sigma^2 + N})\mathbf{1}\mathbf{1}^T) y]_i \\
=& [\sigma^{-2} y - \sigma^{-2}(\frac{\cancel{\sigma^2}}{\sigma^2 + N})\mathbf{1}\mathbf{1}^T) y]_i \\
=& \sigma^{-2} y_i - \frac{\mathbf{1}^T y}{\sigma^2(\sigma^2 + N)}
\end{aligned}$$

$$\begin{aligned}
* : (\mathbf{1}\mathbf{1}^T + \sigma^2 \cdot I)^{-1} \overset{Woodbury}{=} & \sigma^{-2} I - \sigma^{-2} I \cdot \mathbf{1}(I^{-1} + \mathbf{1}^T \sigma^{-2} I \mathbf{1})^{-1} \mathbf{1}^T \sigma^{-2} I \\
=& \sigma^{-2}(I - \sigma^{-2} \cdot \mathbf{1}(I + (\sigma^{-2} \cdot N) I)^{-1} \mathbf{1}^T) \\
=& \sigma^{-2}(I - \sigma^{-2}(1 + \frac{N}{\sigma^2})^{-1} \mathbf{1}\mathbf{1}^T) \\
=& \sigma^{-2}(I - \cancel{\sigma^{-2}}(\frac{\cancel{\sigma^2}}{\sigma^2 + N})\mathbf{1}\mathbf{1}^T) \\
=& \sigma^{-2}(I - (\frac{\mathbf{1}\mathbf{1}^T}{\sigma^2 + N}))
\end{aligned}$$

שאלה 6

$$\begin{aligned}
f_\alpha(x) =& \Sigma_i \alpha_i \overbrace{k(x, x_i)}^{1} \\
=& \Sigma_i [\sigma^{-2} y_i - \frac{\mathbf{1}^T y}{\sigma^2(\sigma^2 + N)}] \\
=& \sigma^{-2} \Sigma_i y_i - N \frac{\mathbf{1}^T y}{\sigma^2(\sigma^2 + N)} \\
=& (\sigma^{-2} - \frac{N}{\sigma^2(\sigma^2 + N)}) \mathbf{1}^T y \\
=& (\frac{\cancel{\sigma^2} + \cancel{N - N}}{\cancel{\sigma^2}(\sigma^2 + N)}) \mathbf{1}^T y \\
=& \frac{1}{\sigma^2 + N} \mathbf{1}^T y
\end{aligned}$$

לכן:

$$f_\alpha(x) \underset{\sigma^2\to 0 \wedge N\to\infty}{\to} 0$$

נקבל:

$$\begin{aligned} Train\ error =& \frac{1}{N}\Sigma_i||sign(x_i) - f_\alpha(x_i)||^2 \\ \underset{\sigma^2\to 0\wedge N\to\infty}{=}& \frac{1}{N}\Sigma_i||sign(x_i) - 0||^2 \\ =& \frac{1}{N}\Sigma_i 1 = 1 \end{aligned}$$

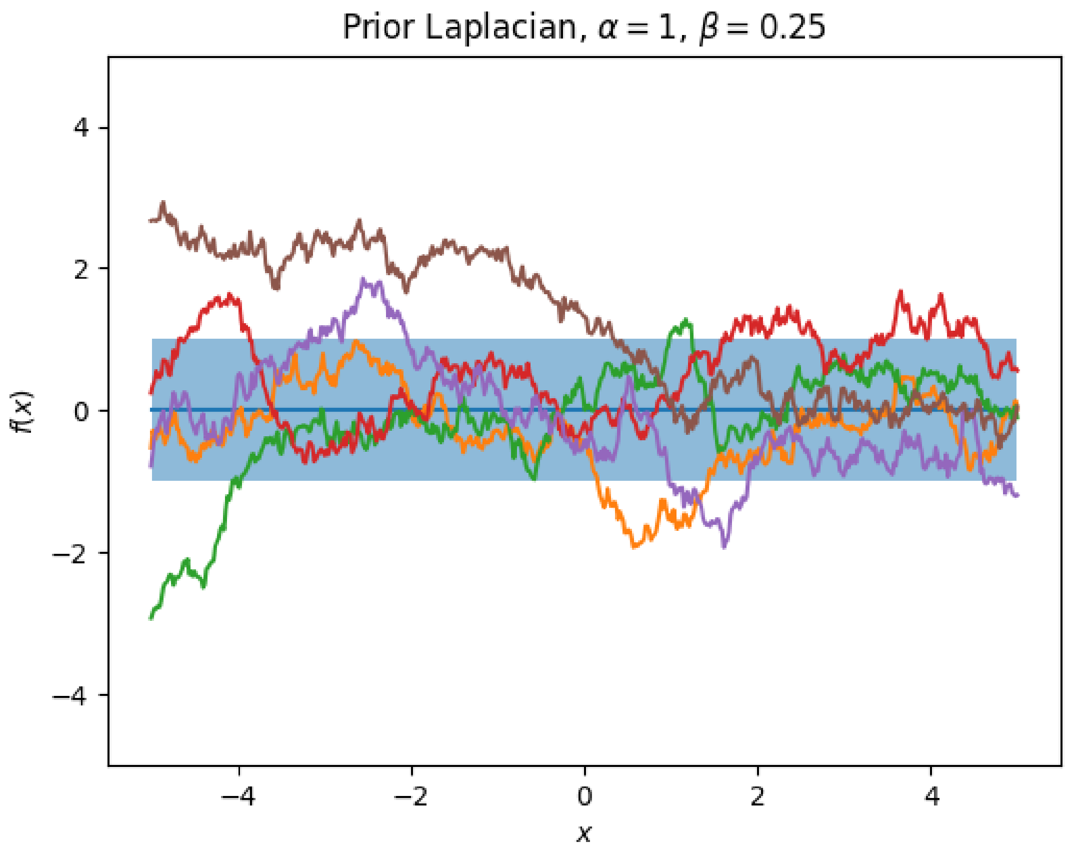
שאלה 7

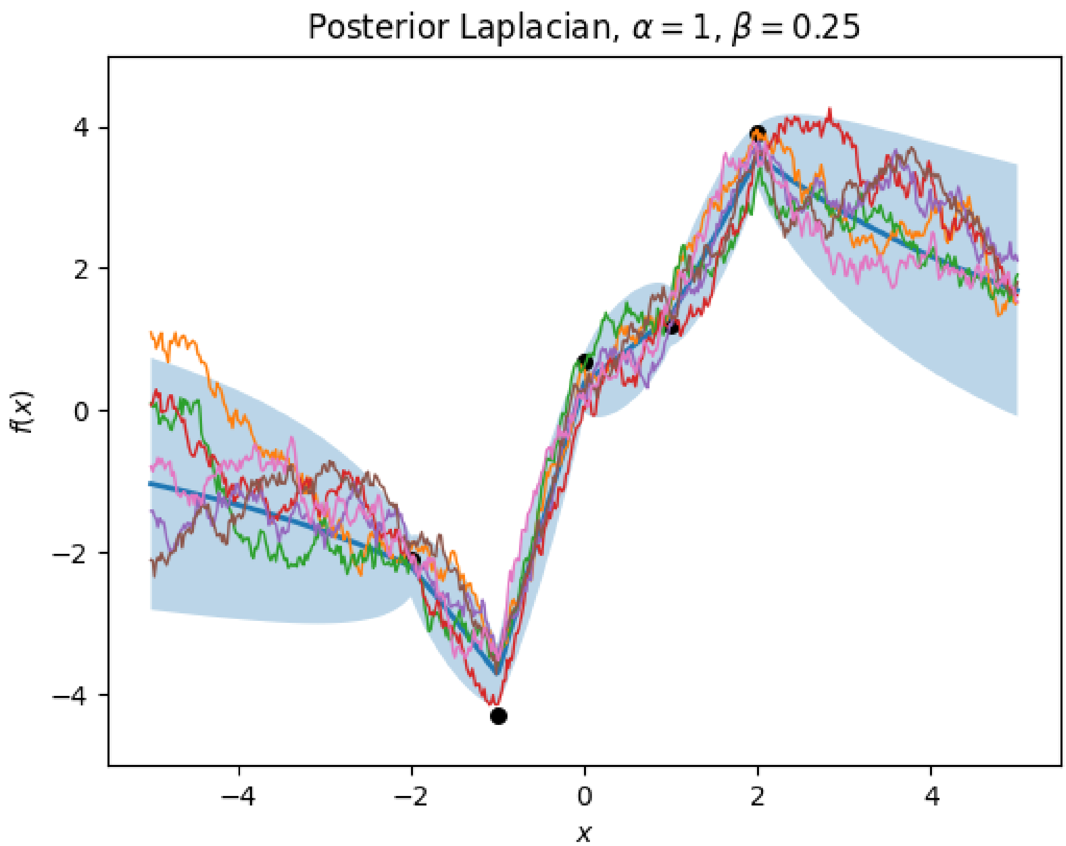
ראינו בתרגול ש-RBF הוא $kernel\ PD$ לכן כאשר $\sigma^2 \to 0$ אז שגיאת האימון היא 0.

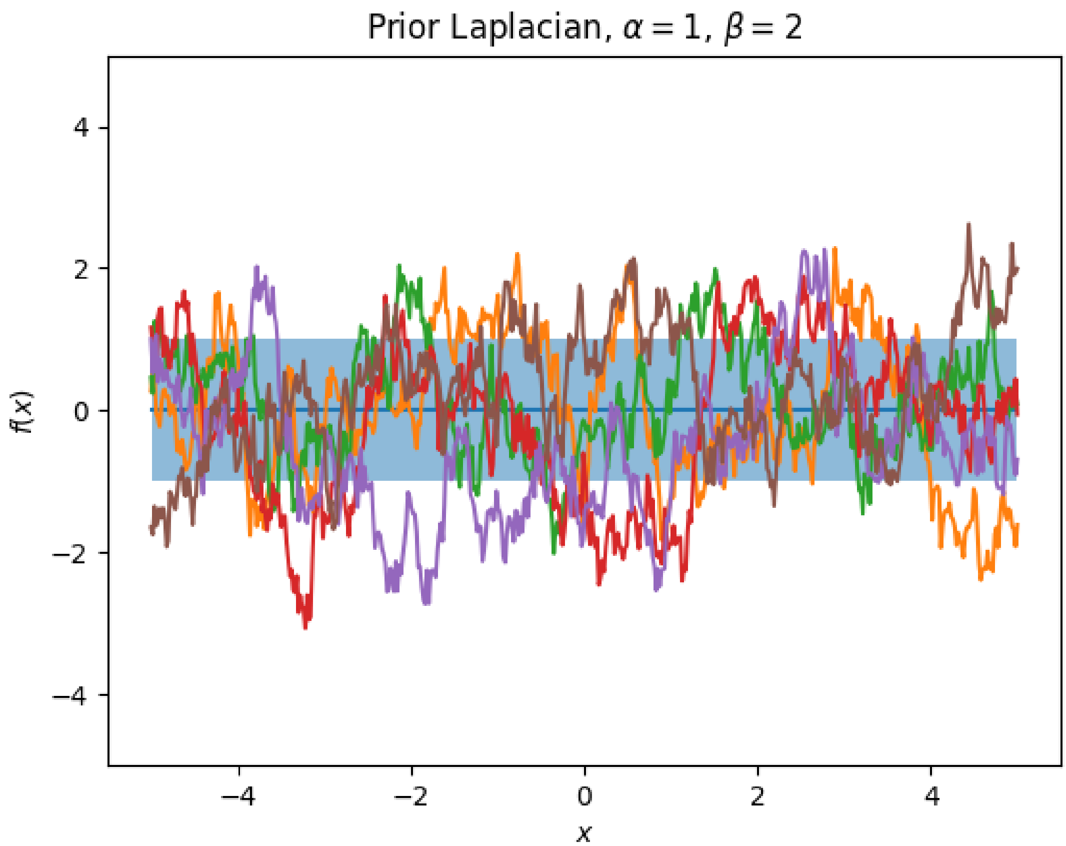
פרקטי

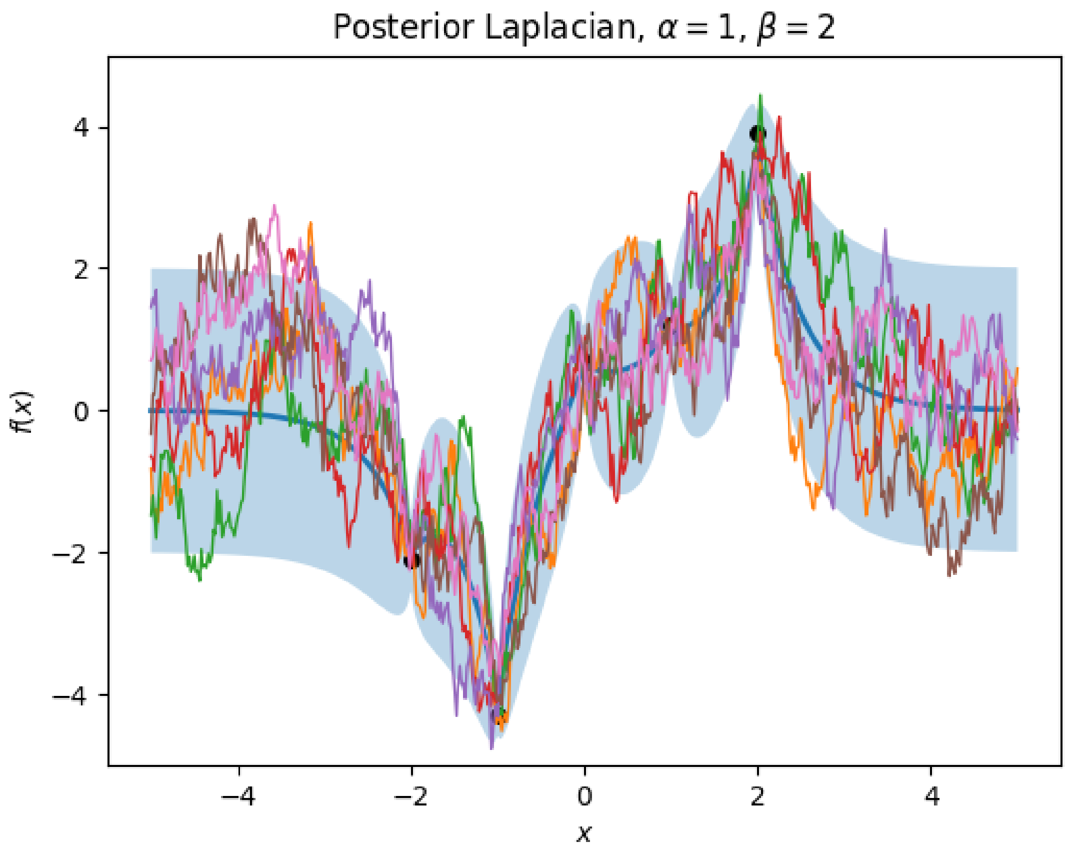
שאלות 2,3

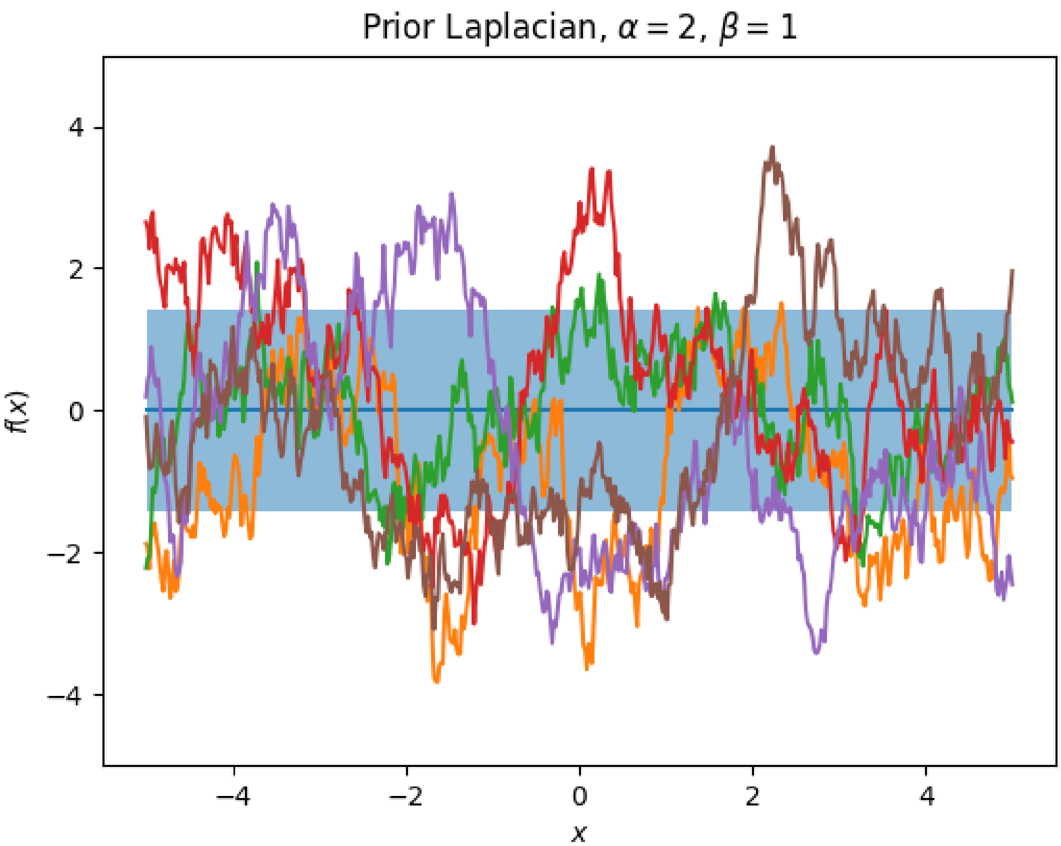
$Laplacian$:


Prior Laplacian, $\alpha = 1$, $\beta = 0.25$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

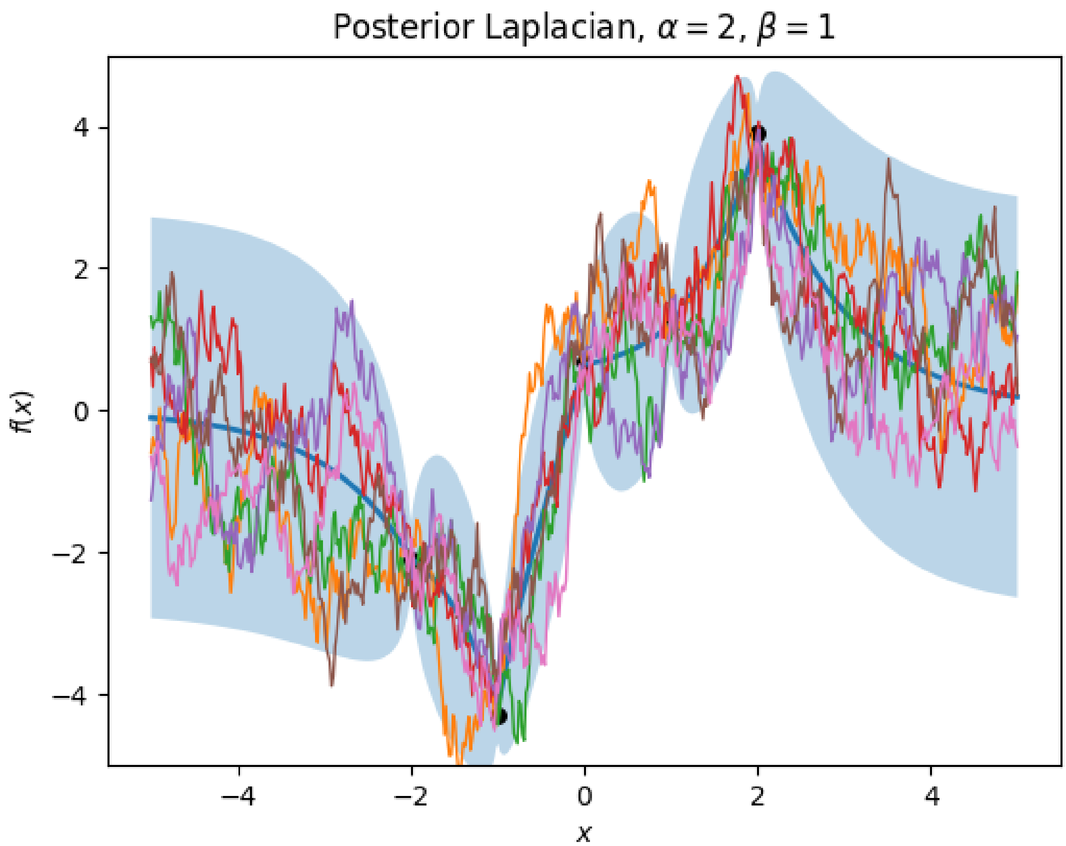

Posterior Laplacian, $\alpha = 1$, $\beta = 0.25$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x


Prior Laplacian, $\alpha = 1$, $\beta = 2$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4

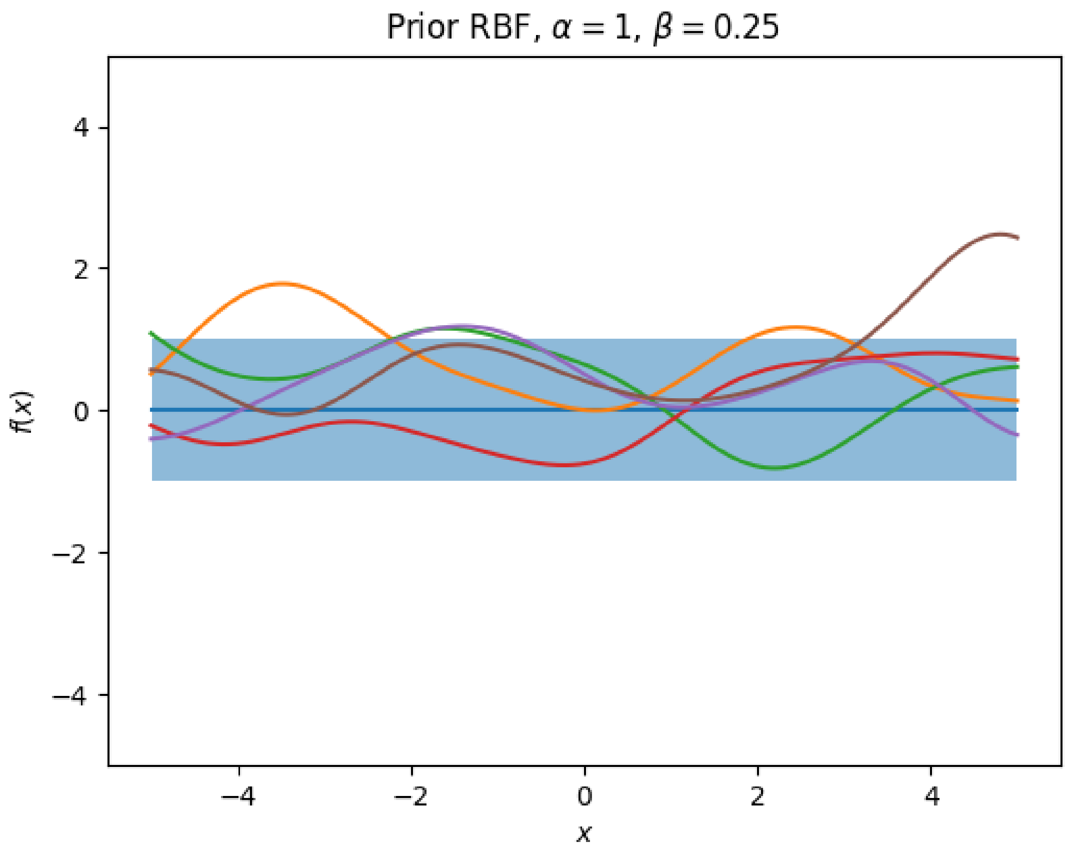

Posterior Laplacian, $\alpha = 1$, $\beta = 2$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x



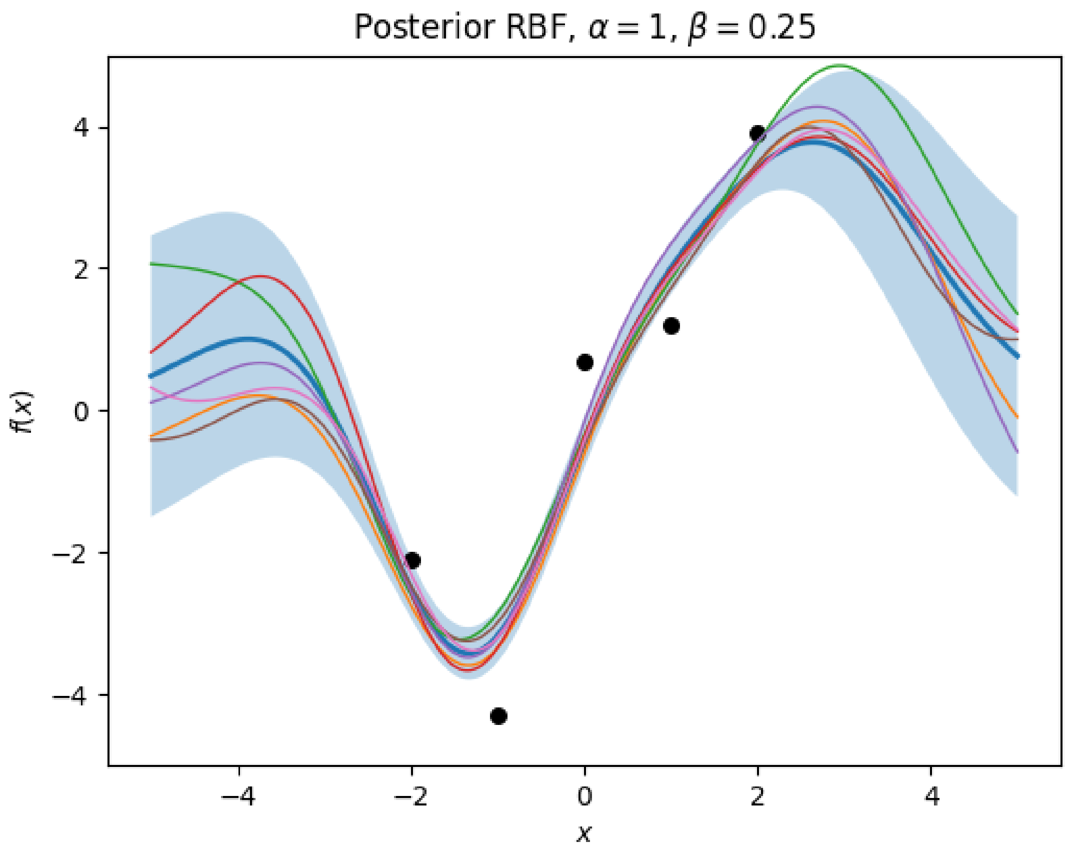
Prior Laplacian, $\alpha = 2$, $\beta = 1$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4



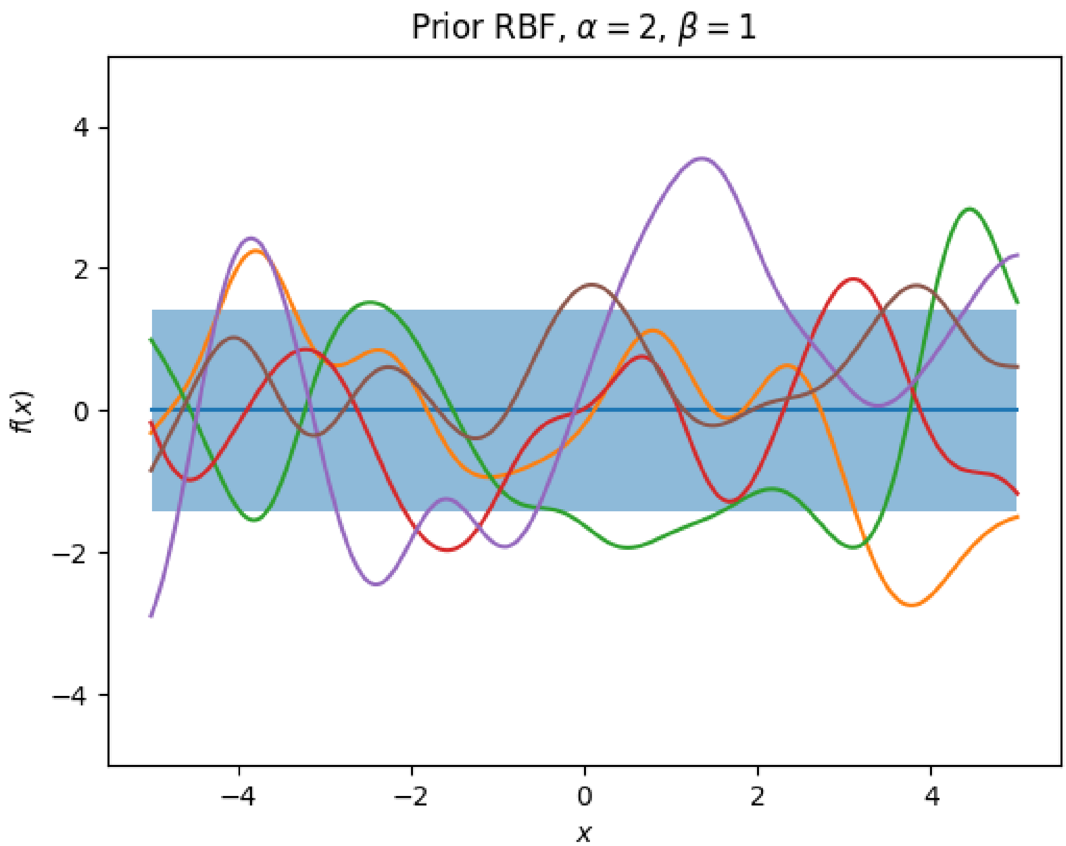
$:RBF$

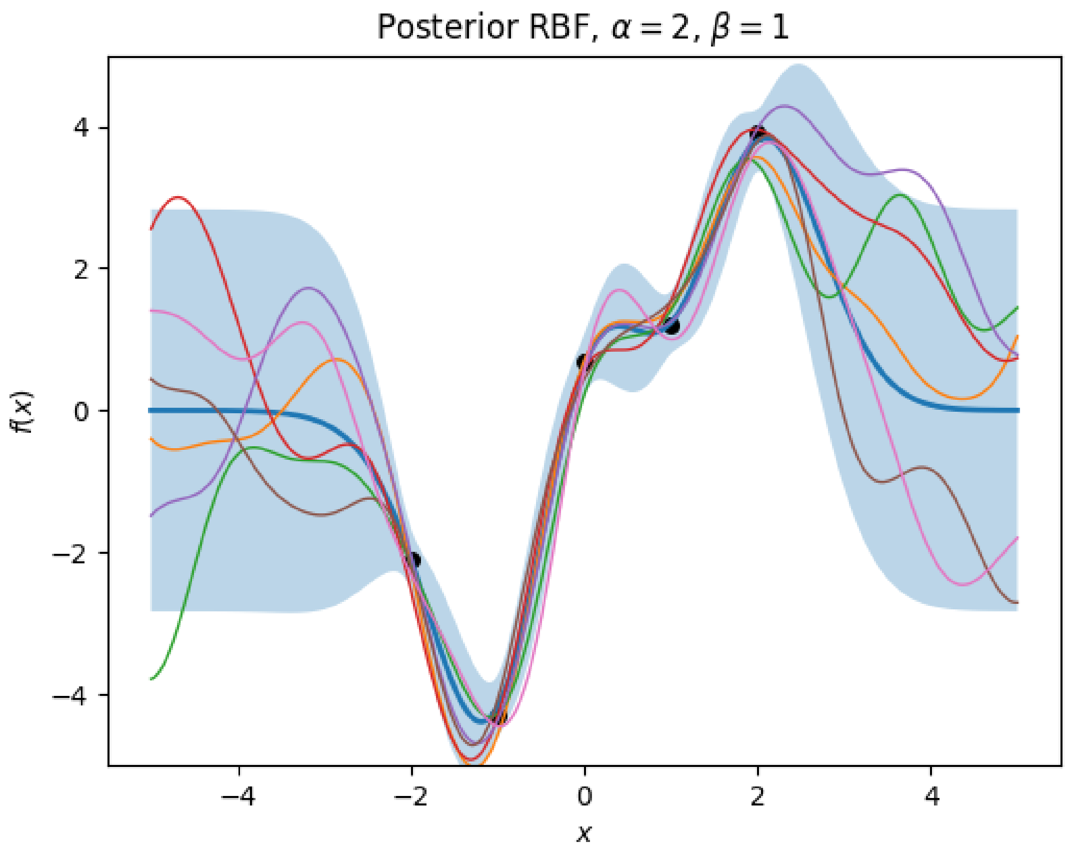


Prior RBF, $\alpha = 1$, $\beta = 0.25$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

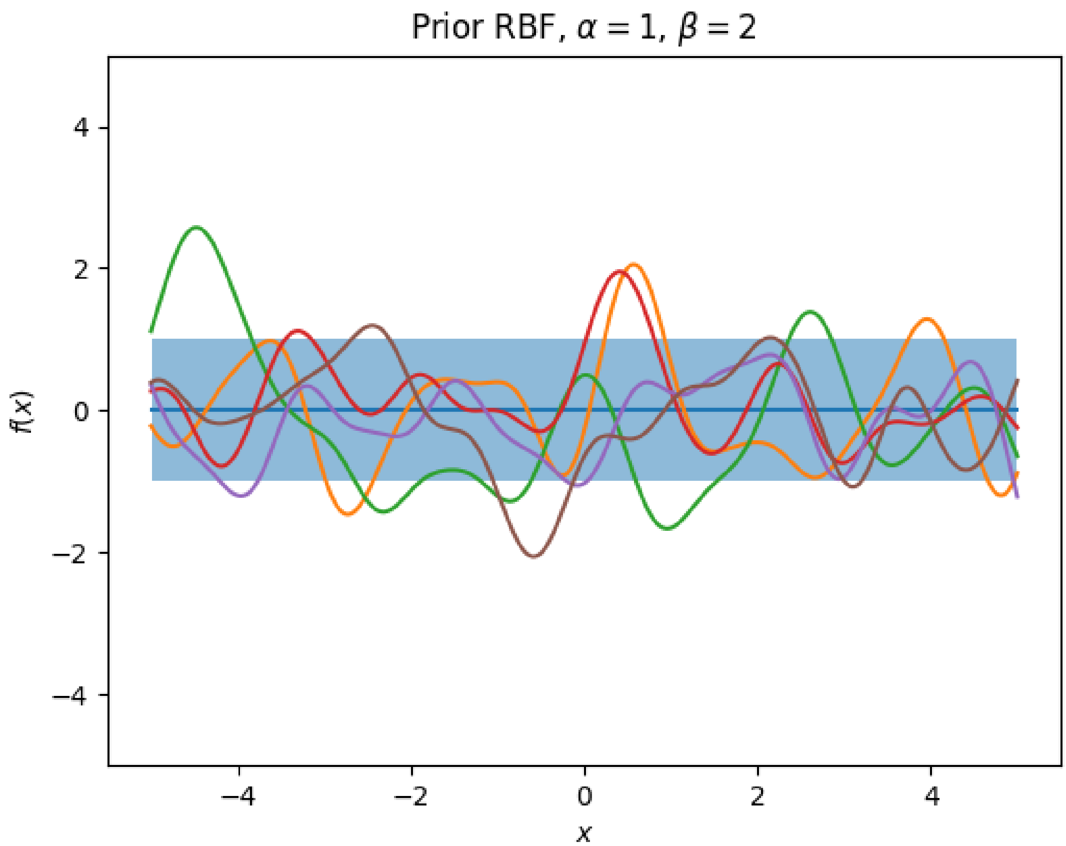


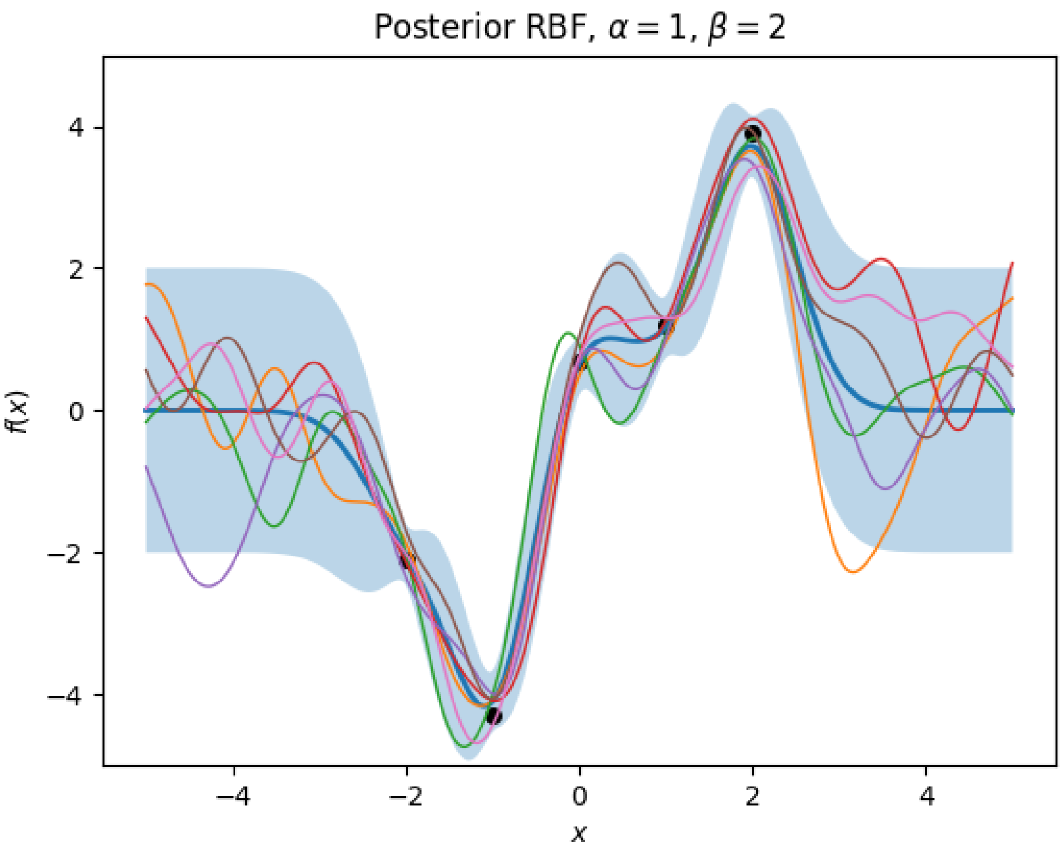
Posterior RBF, $\alpha = 1$, $\beta = 0.25$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4


Prior RBF, $\alpha = 2$, $\beta = 1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

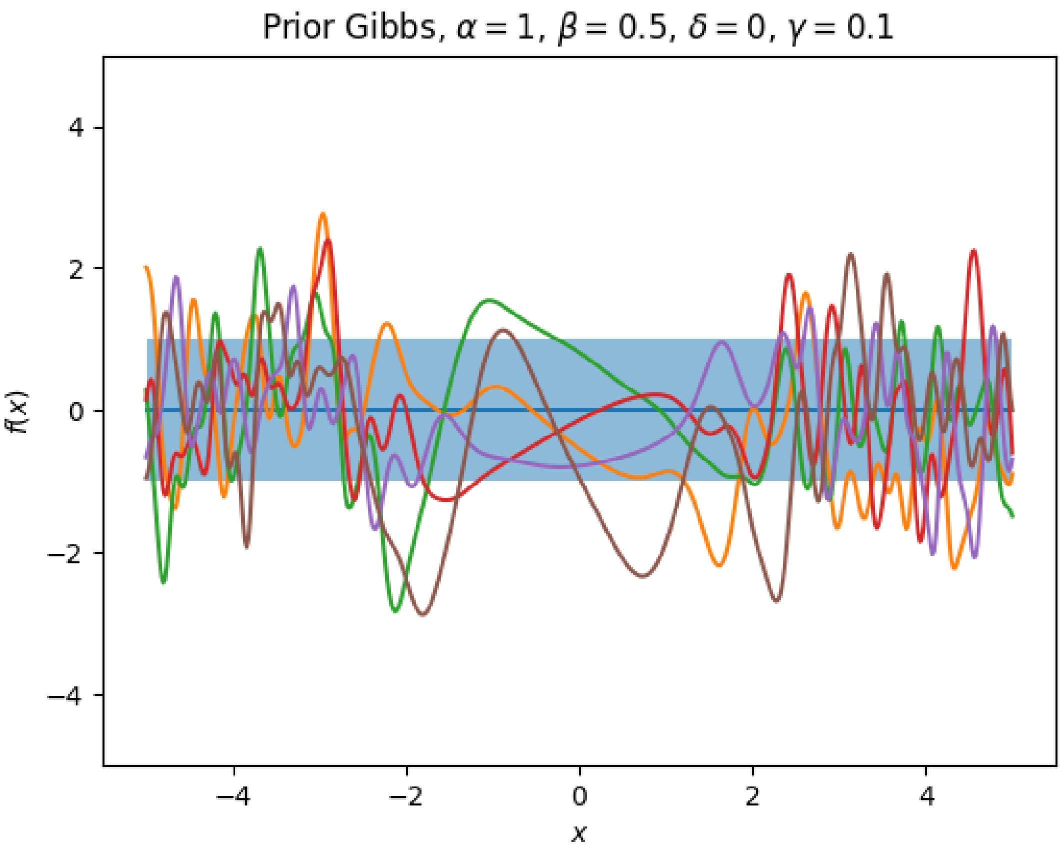


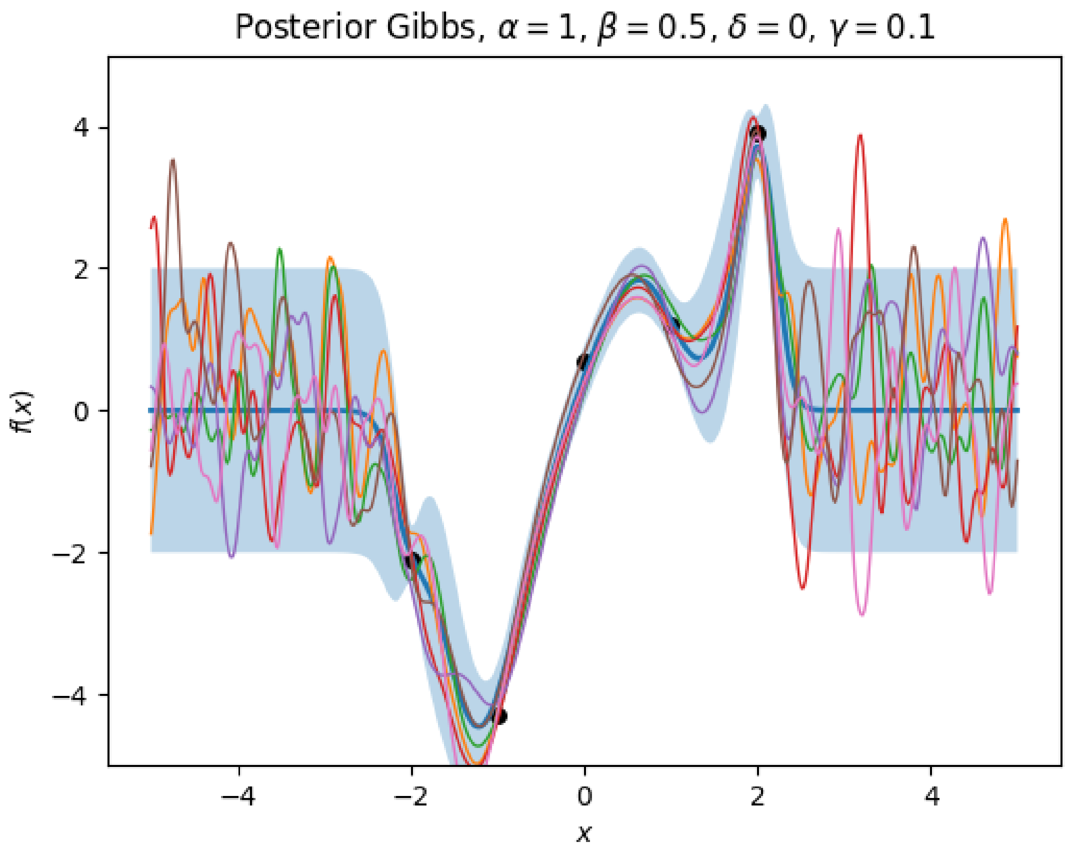
Posterior RBF, $\alpha = 2$, $\beta = 1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x


Prior RBF, $\alpha = 1$, $\beta = 2$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4

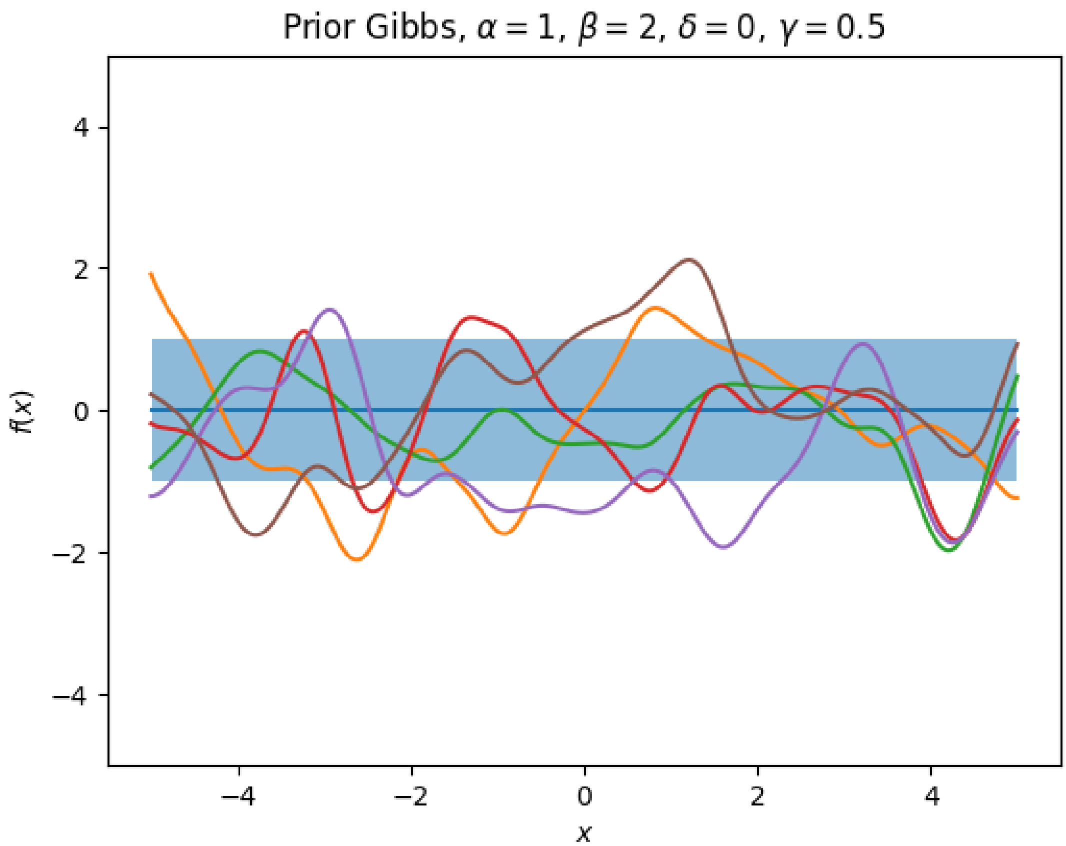


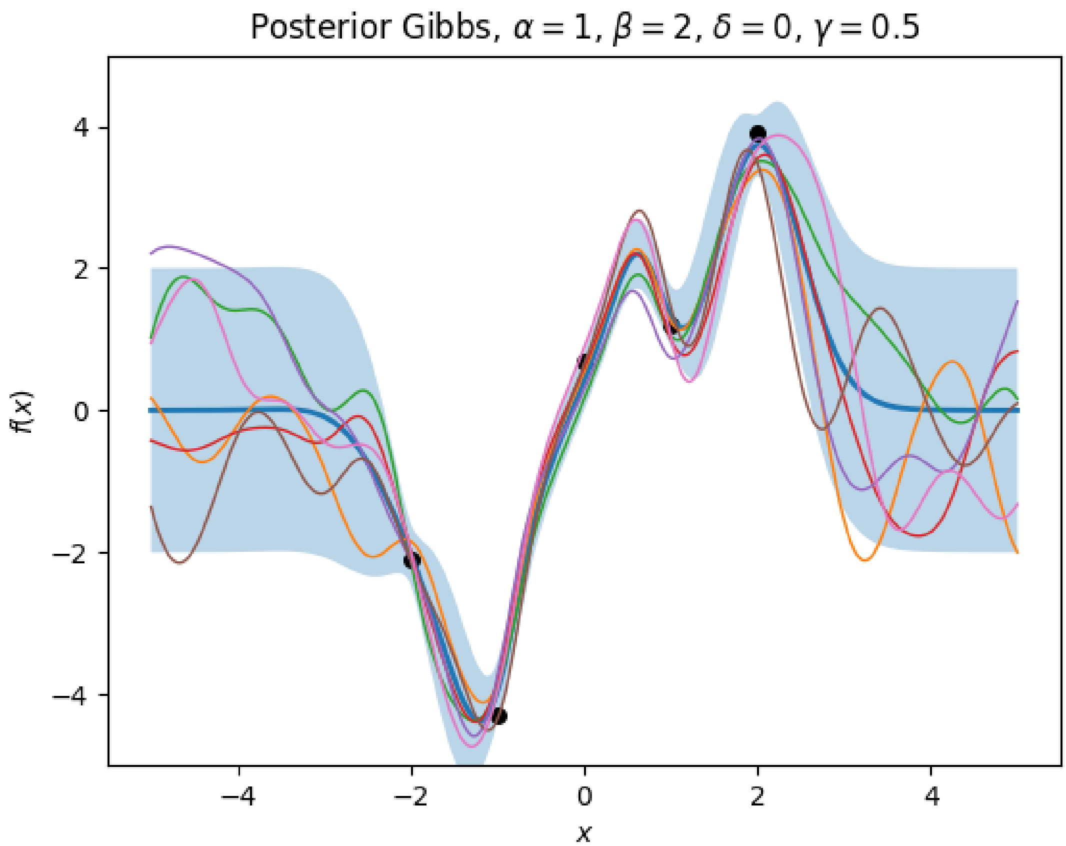
$: Gibbs$


Prior Gibbs, $\alpha = 1$, $\beta = 0.5$, $\delta = 0$, $\gamma = 0.1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

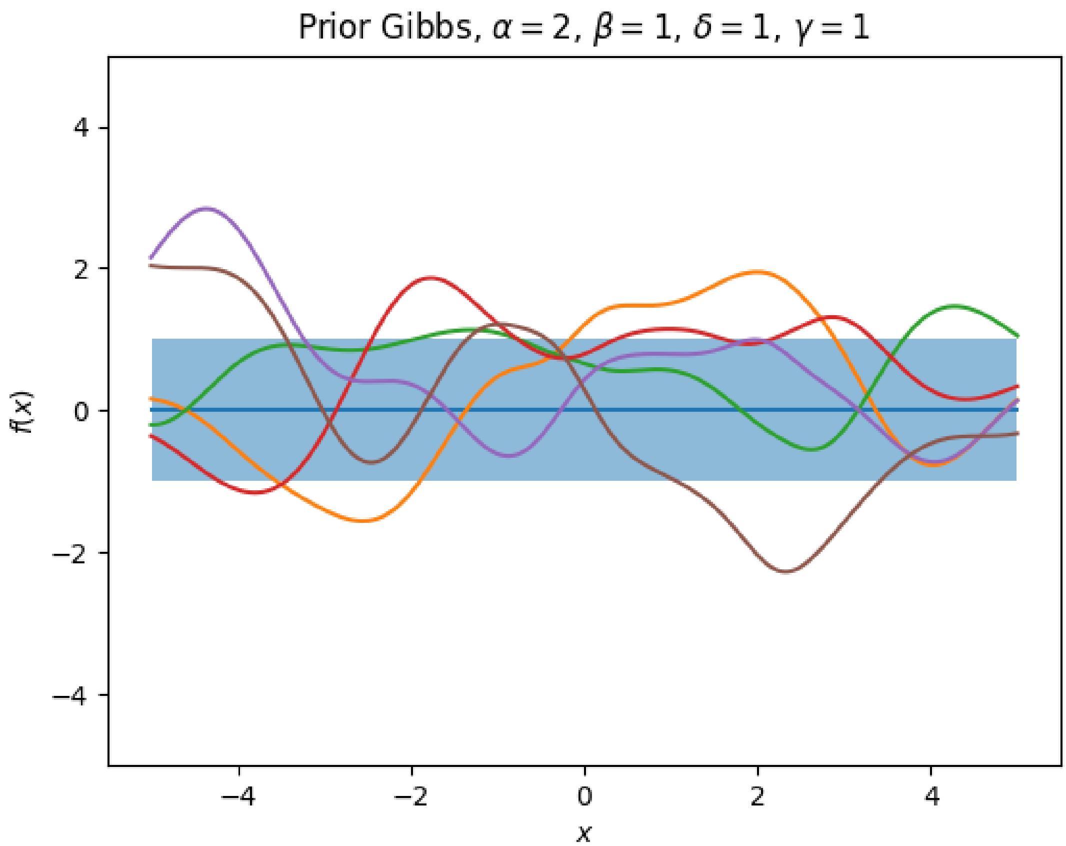


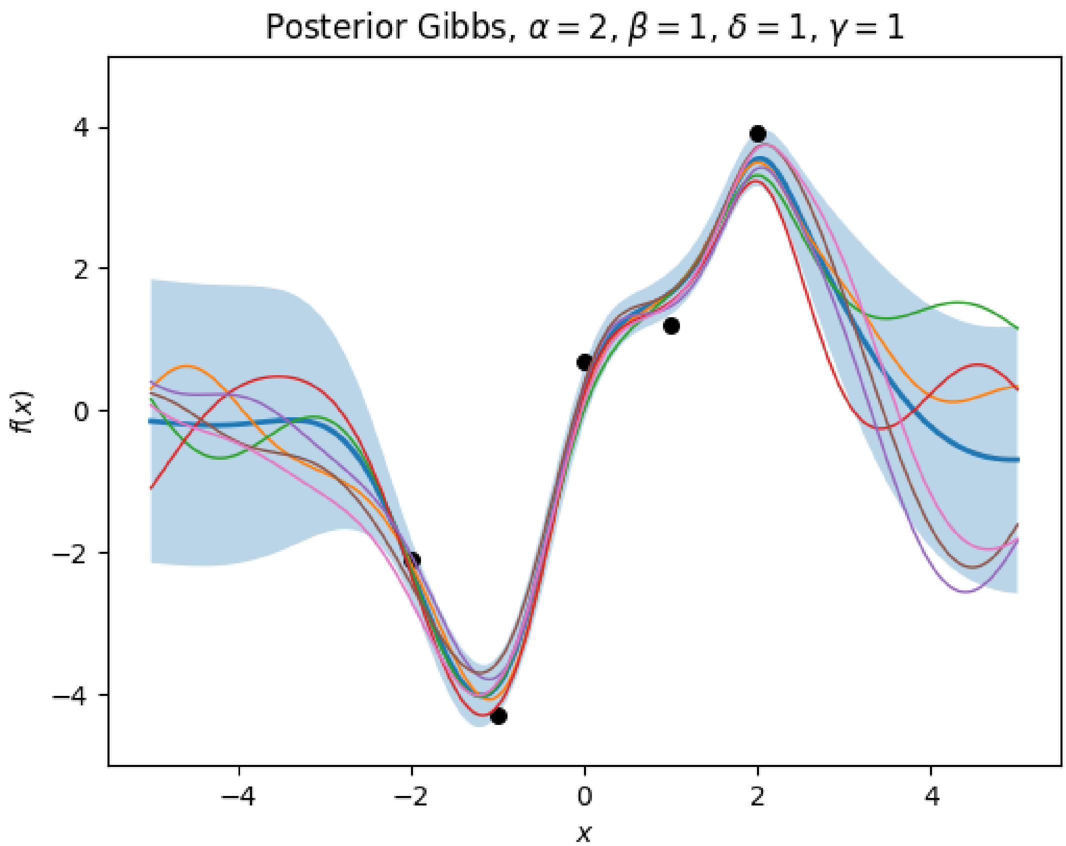
Posterior Gibbs, $\alpha = 1$, $\beta = 0.5$, $\delta = 0$, $\gamma = 0.1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x


Prior Gibbs, $\alpha = 1$, $\beta = 2$, $\delta = 0$, $\gamma = 0.5$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4

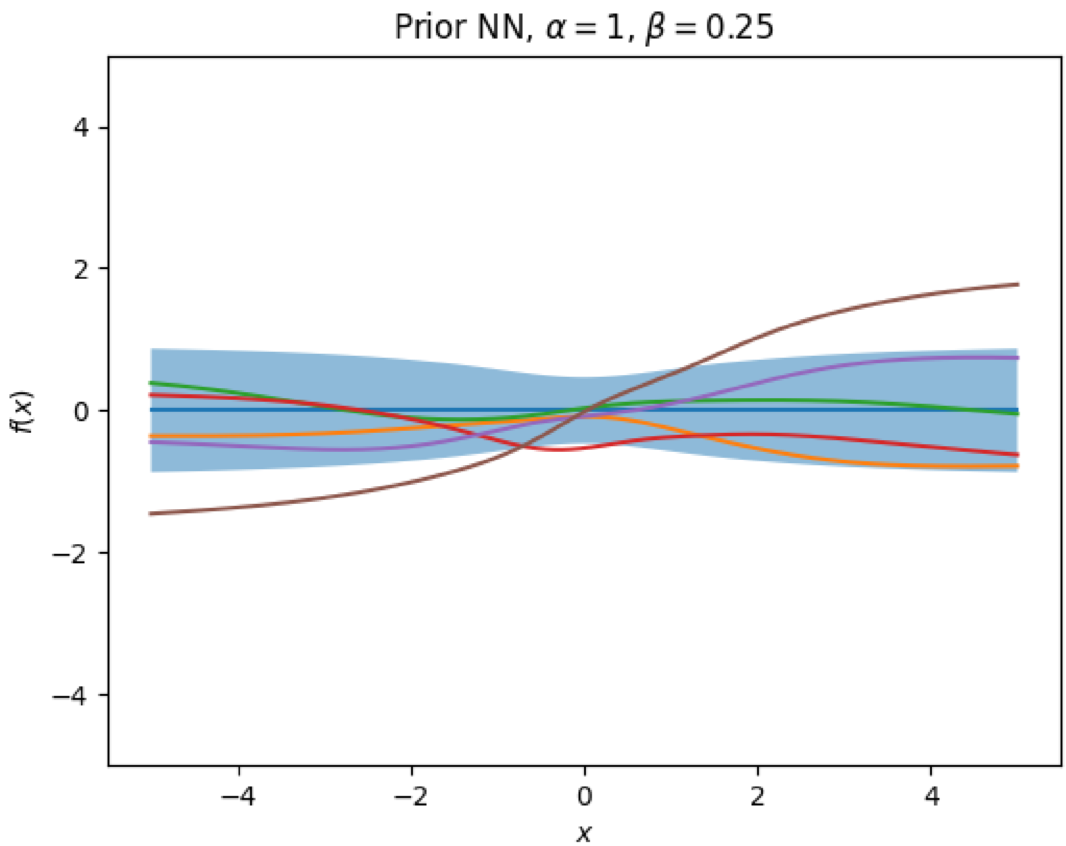


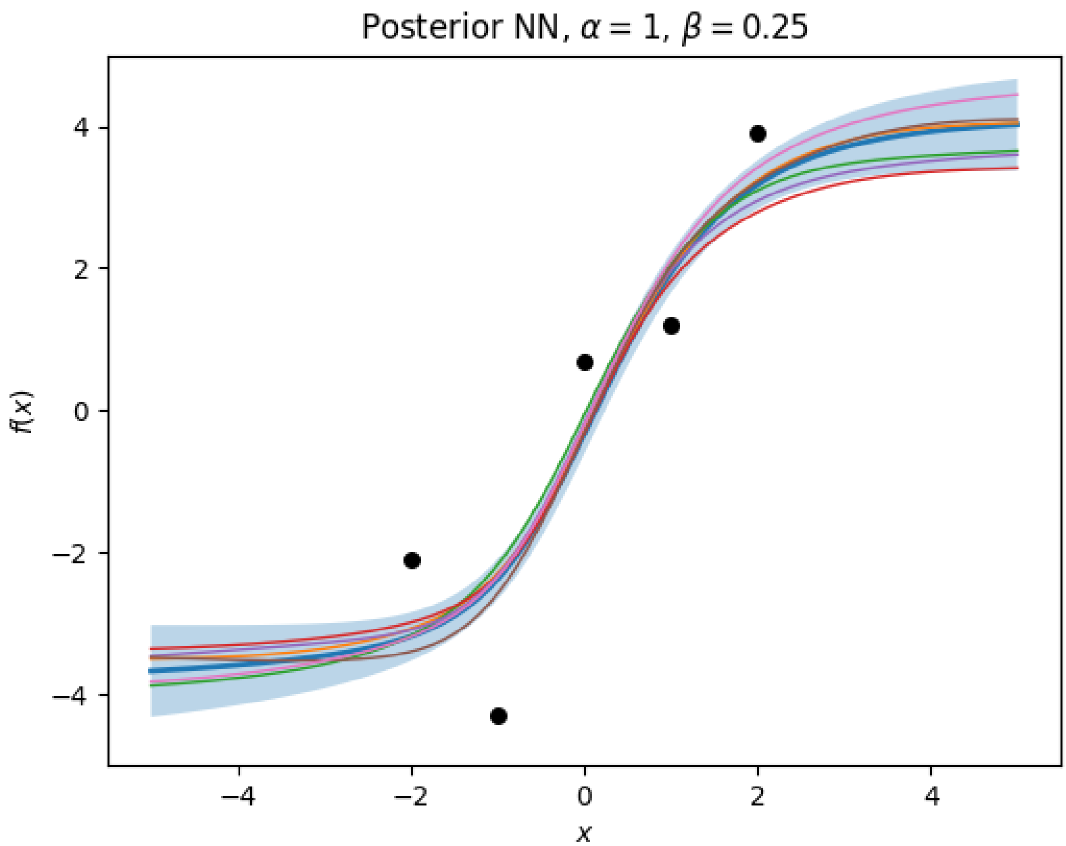
Posterior Gibbs, $\alpha = 1$, $\beta = 2$, $\delta = 0$, $\gamma = 0.5$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x


Prior Gibbs, $\alpha = 2$, $\beta = 1$, $\delta = 1$, $\gamma = 1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

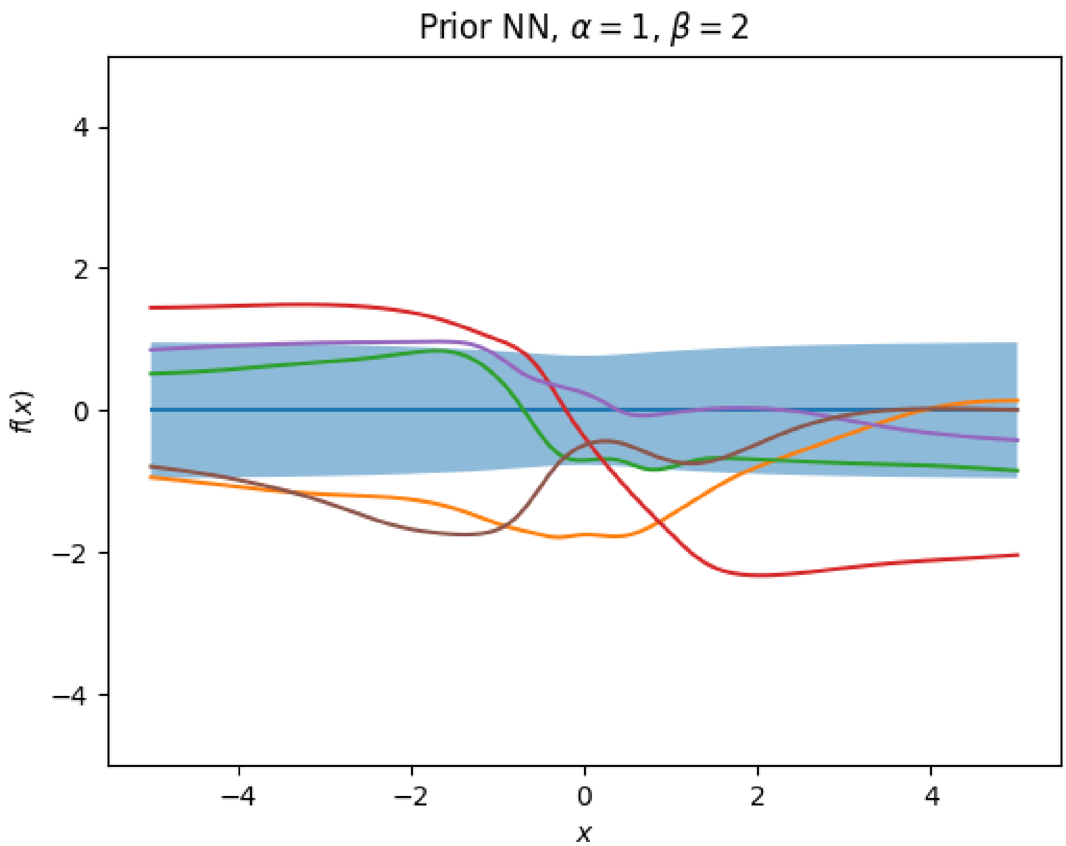

Posterior Gibbs, $\alpha = 2$, $\beta = 1$, $\delta = 1$, $\gamma = 1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

$:NN$

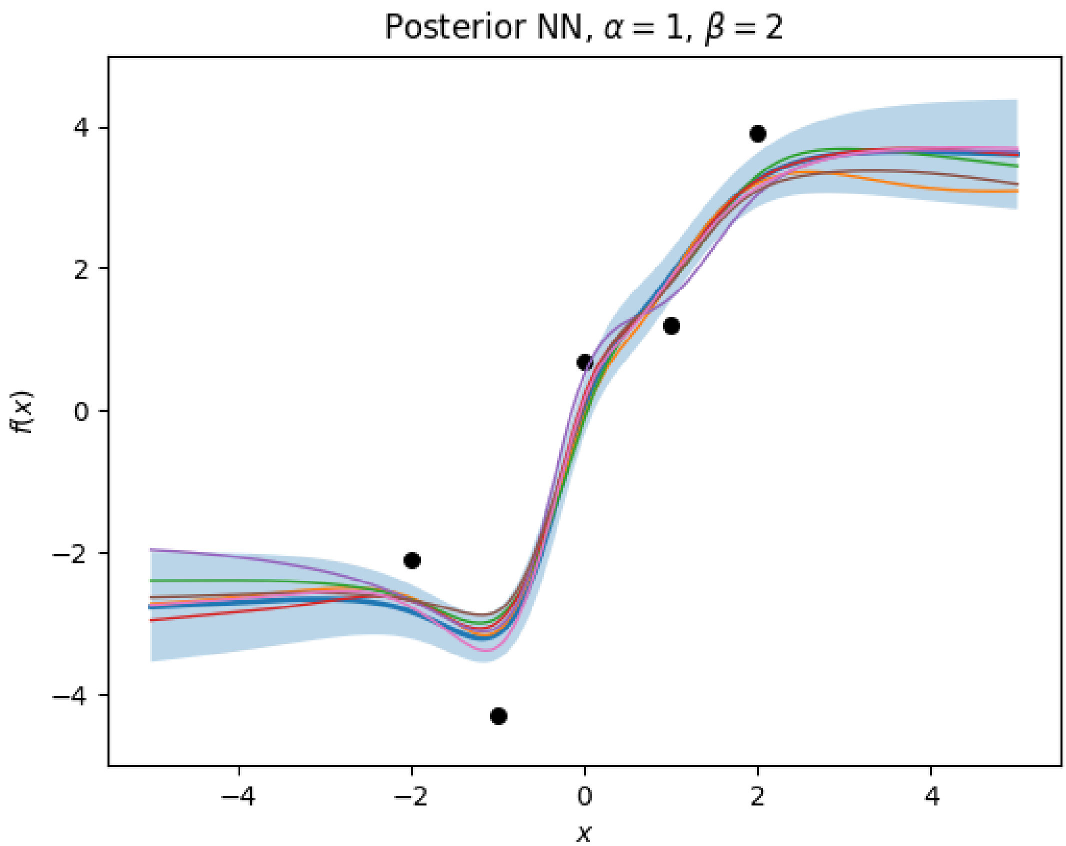

Prior NN, $\alpha = 1$, $\beta = 0.25$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x



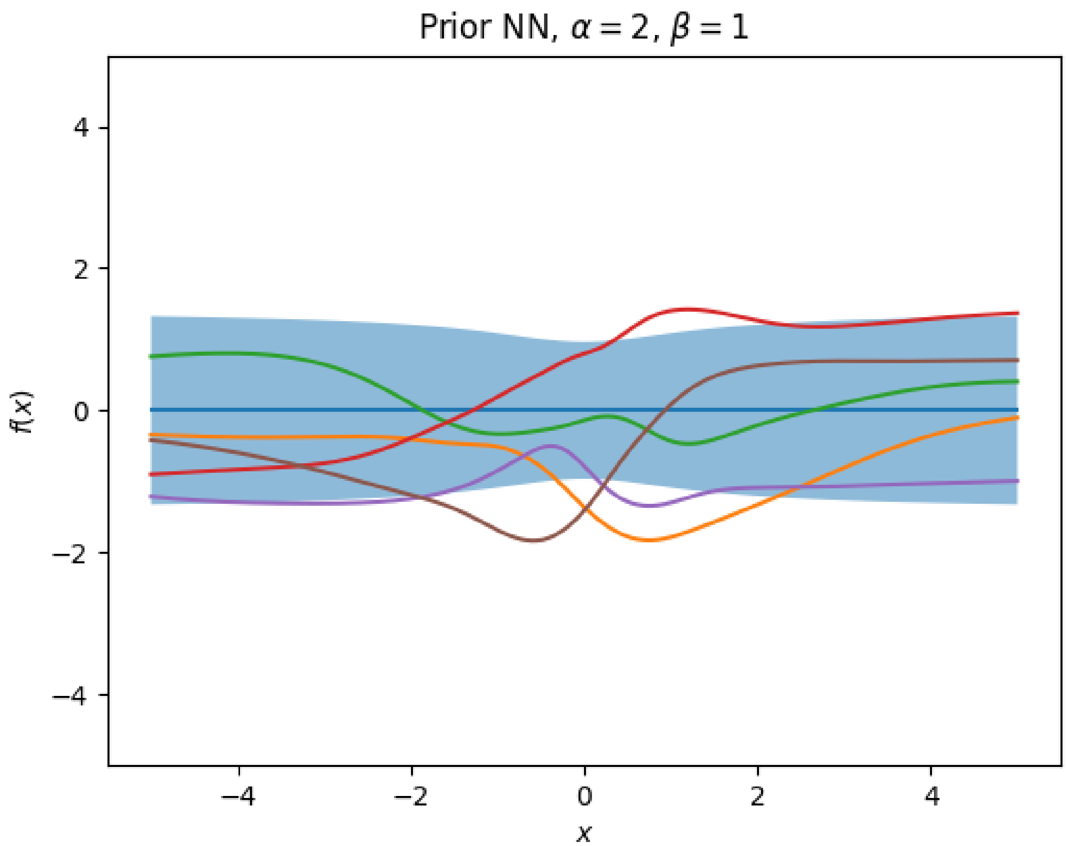
Posterior NN, $\alpha = 1$, $\beta = 0.25$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4

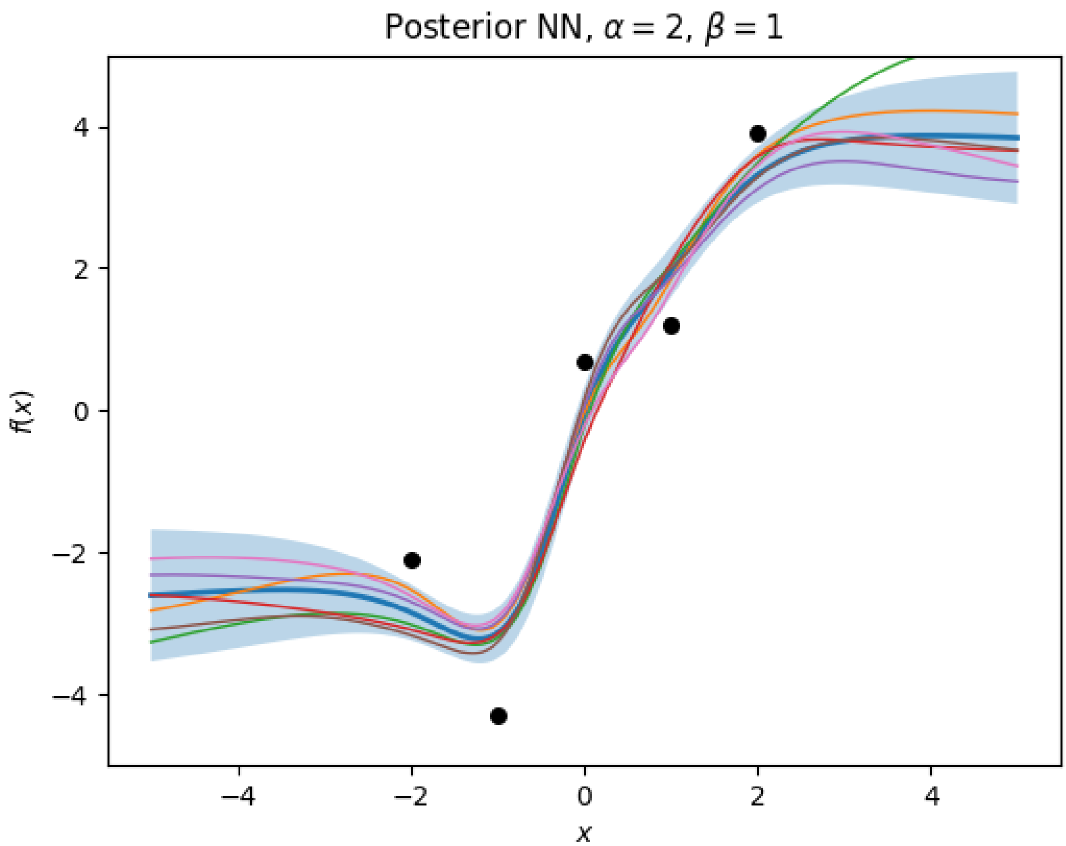


Prior NN, $\alpha = 1$, $\beta = 2$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x



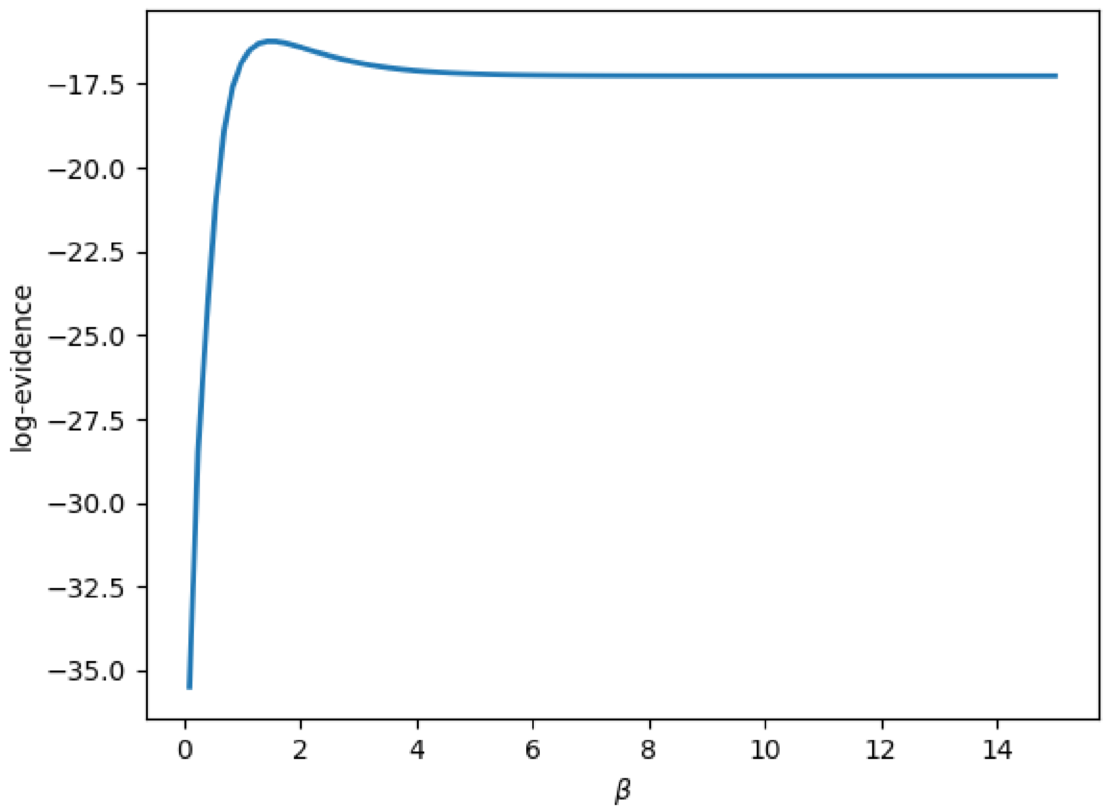
Posterior NN, $\alpha = 1$, $\beta = 2$
$f(x)$
x
4
2
0
−2
−4
−4
−2
0
2
4


Prior NN, $\alpha = 2$, $\beta = 1$
4
2
0
−2
−4
$f(x)$
−4
−2
0
2
4
x

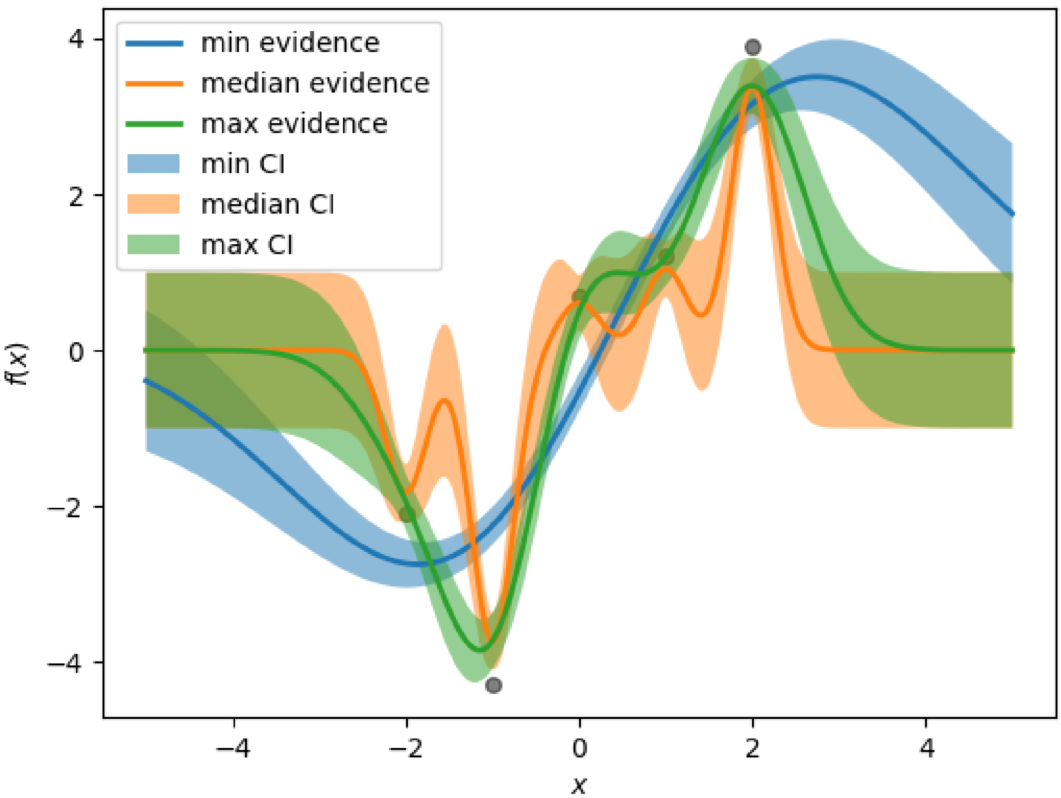


Posterior NN, $\alpha = 2$, $\beta = 1$
$f(x)$
x
4
2
0
−2
−4

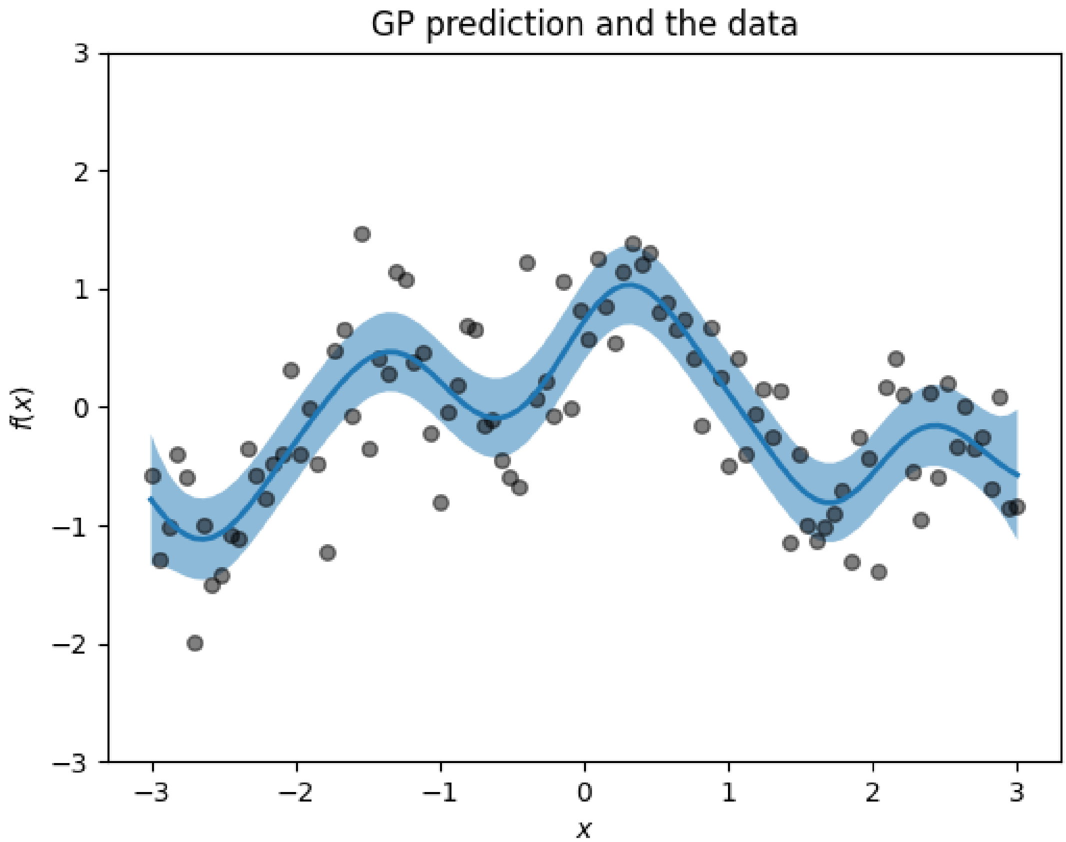
שאלה 4



עבור $\beta = 1.441$ יש את הערך $evidence$ הטוב ביותר.

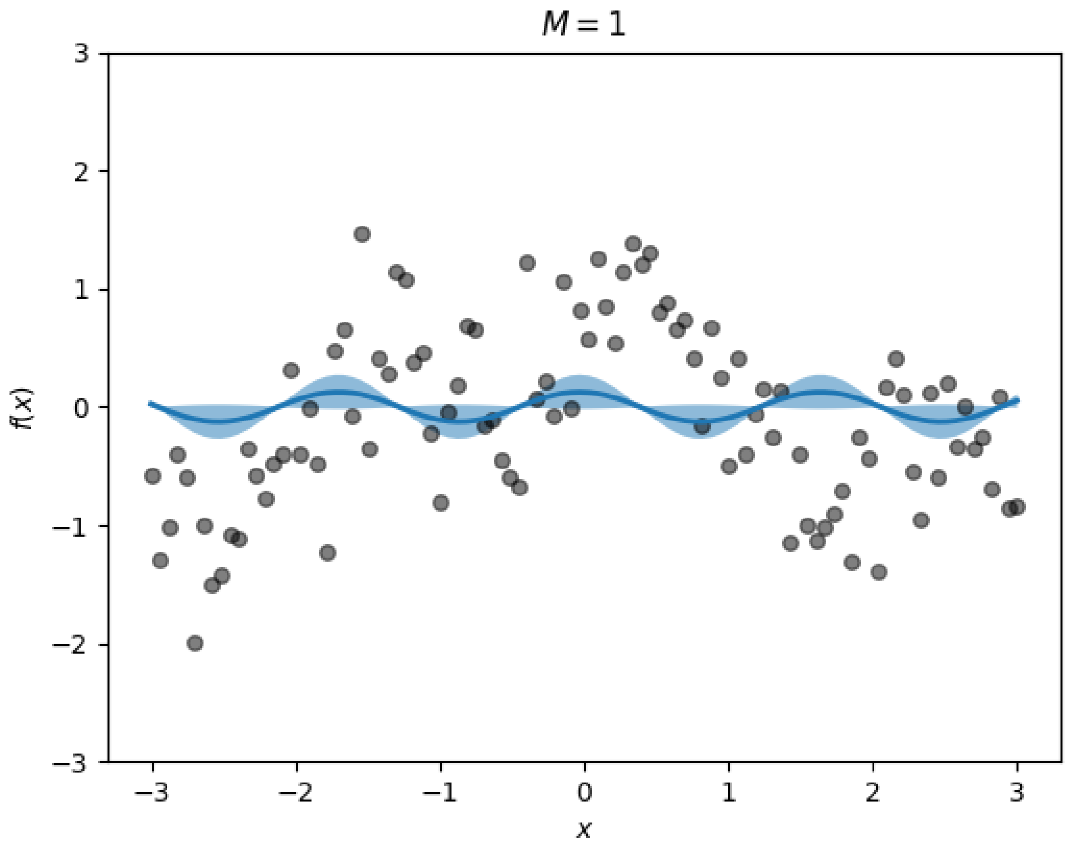

min evidence
median evidence
max evidence
min CI
median CI
max CI
4
2
0
−2
−4
f(x)
−4
−2
0
2
4
x

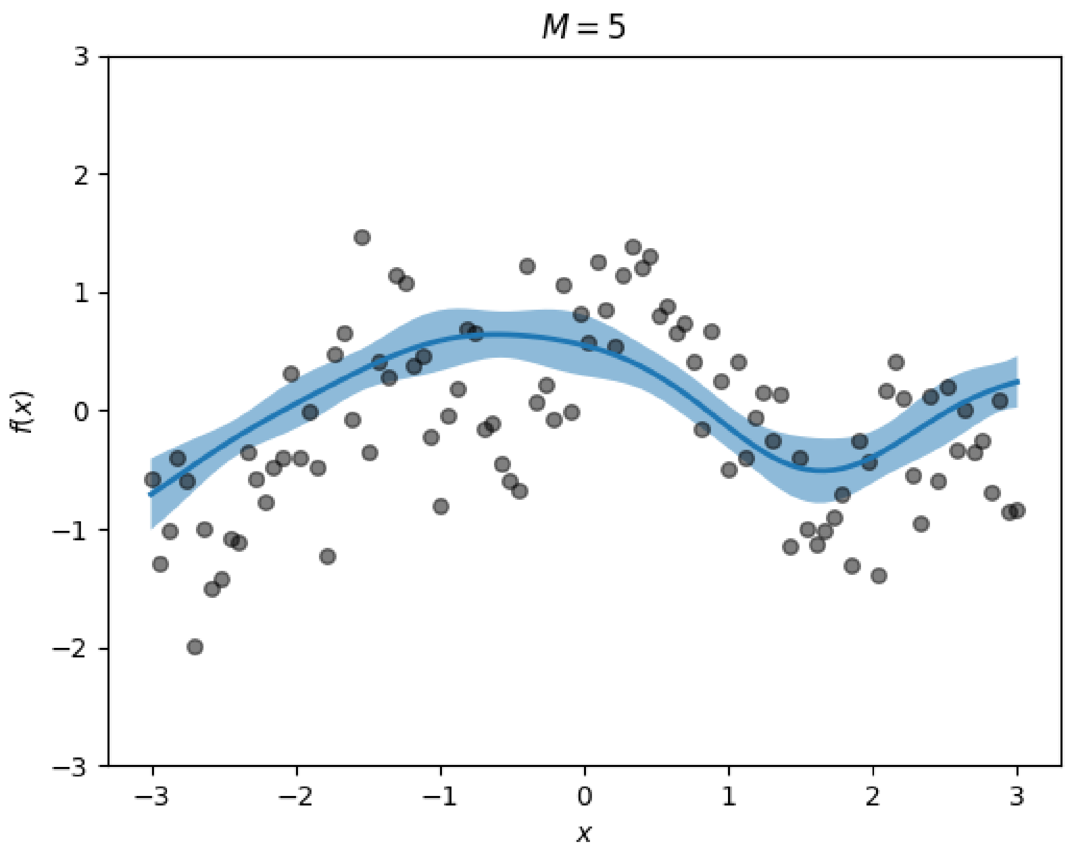
שאלה 5



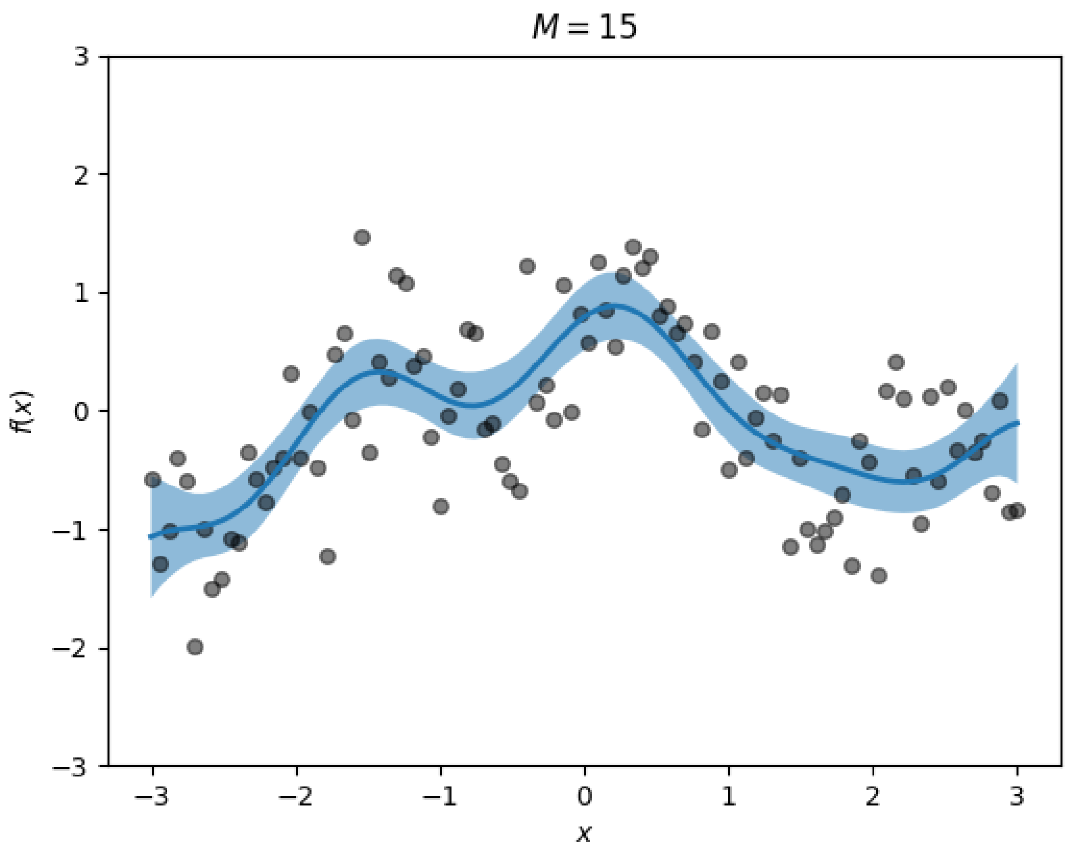
וקיבלנו $Average\ squared\ error\ of\ the\ GP\ is: 0.20$

שאלה 7

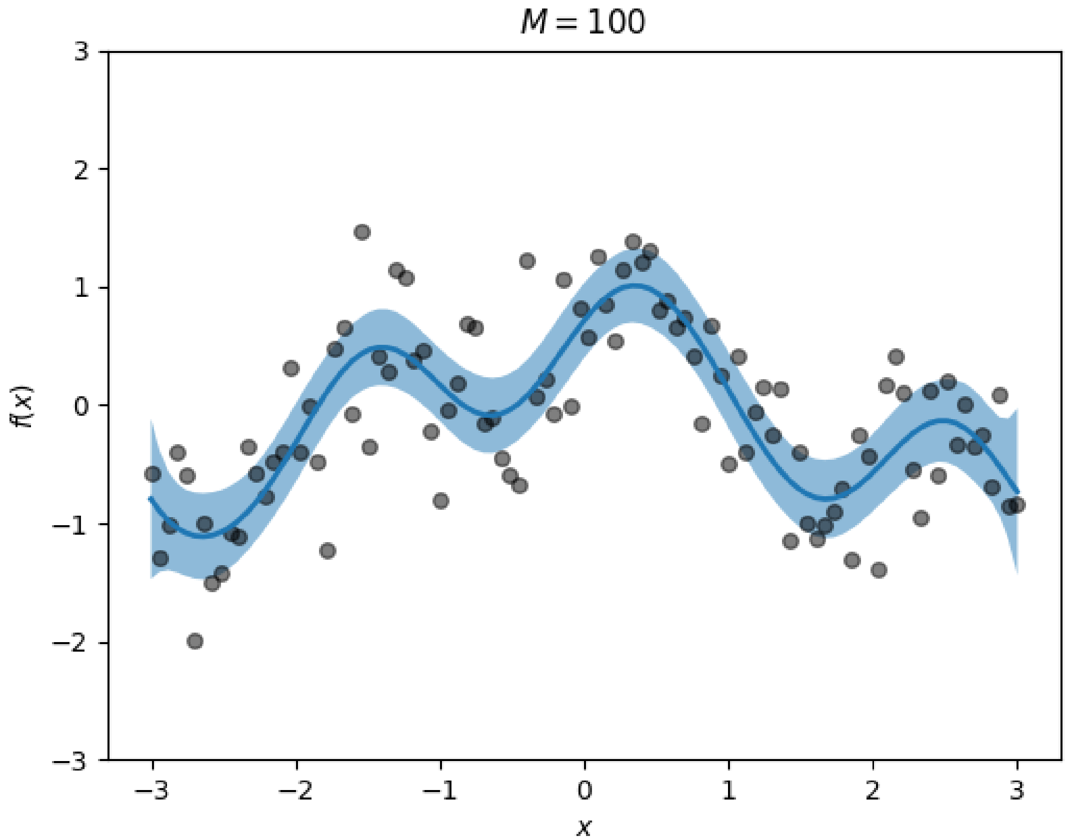


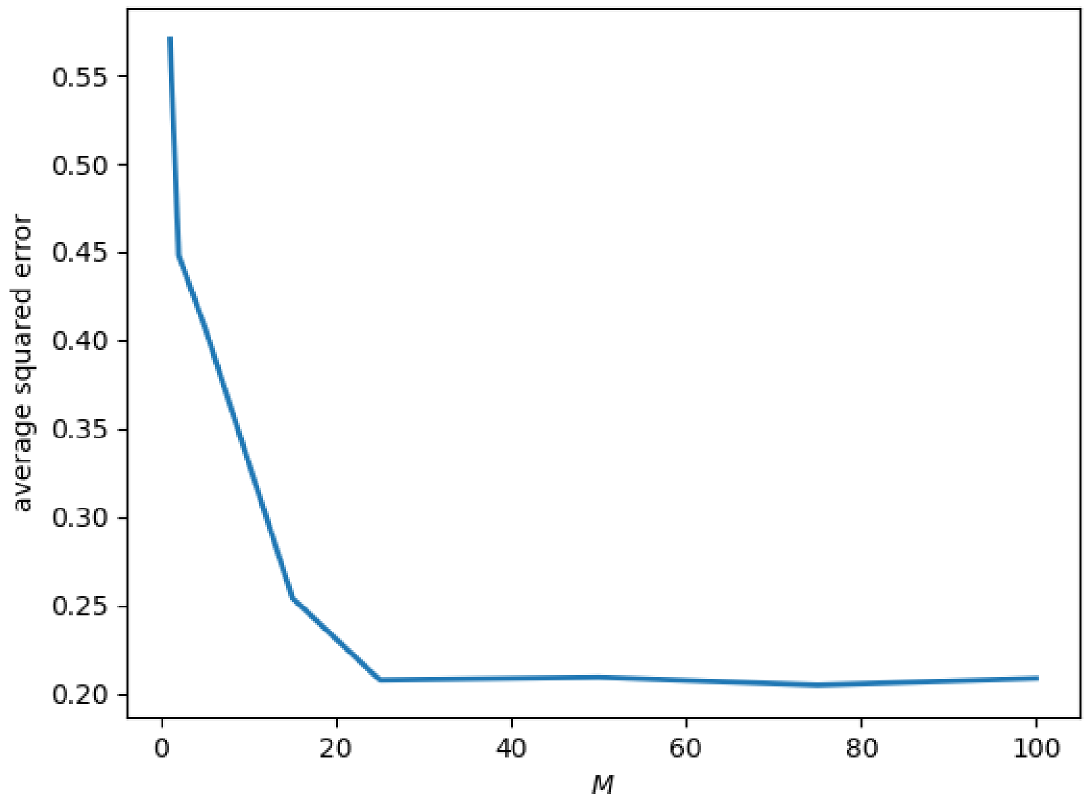


$M = 5$
3
2
1
0
−1
−2
−3
$f(x)$
−3
−2
−1
0
1
2
3
x



$M = 15$
3
2
1
0
−1
−2
−3
$f(x)$
−3
−2
−1
0
1
2
3
x


$M = 100$
3
2
1
0
−1
−2
−3
$f(x)$
−3
−2
−1
0
1
2
3
x




```
Average squared error with M=1 random features: 0.57
Average squared error with M=2 random features: 0.45
Average squared error with M=5 random features: 0.41
Average squared error with M=15 random features: 0.25
Average squared error with M=25 random features: 0.21
Average squared error with M=50 random features: 0.21
Average squared error with M=75 random features: 0.20
Average squared error with M=100 random features: 0.21
```