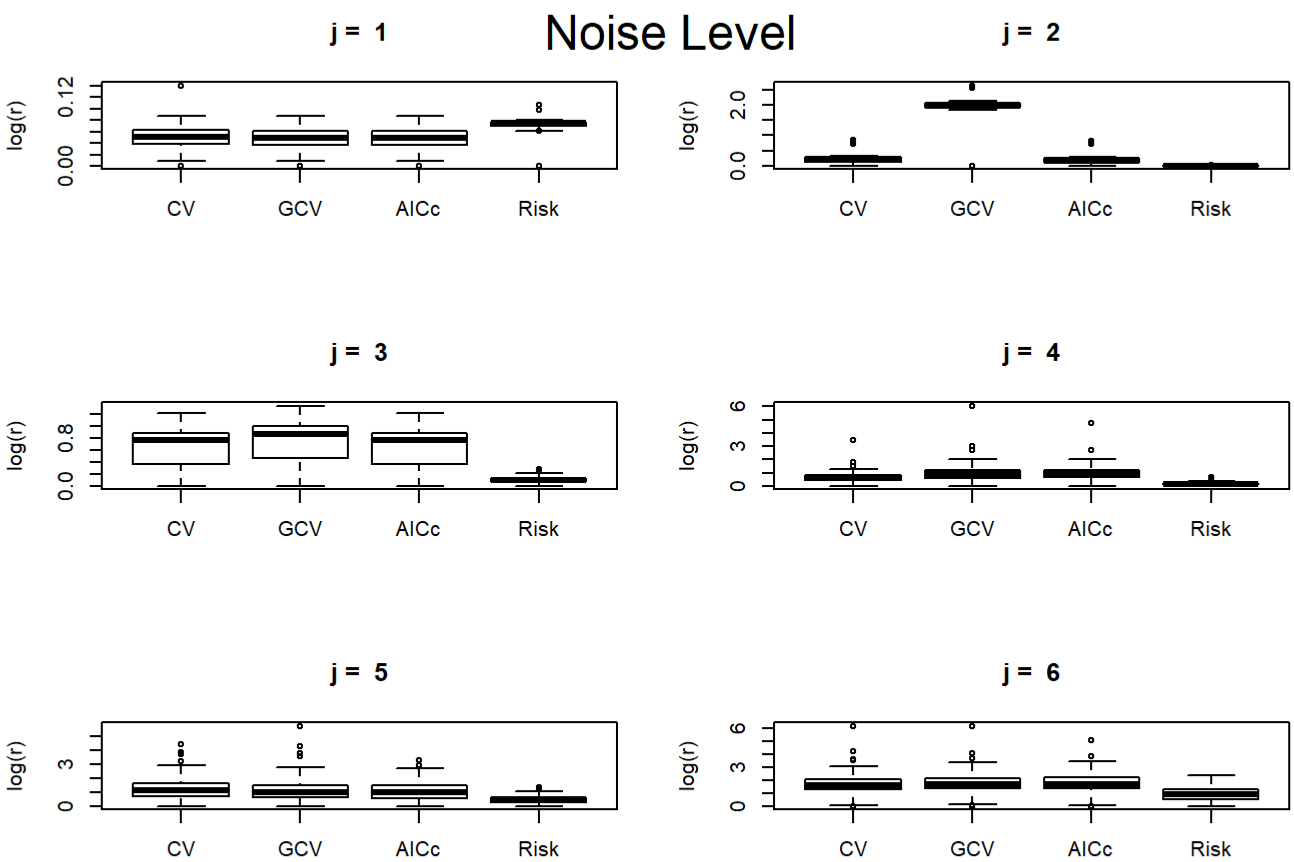
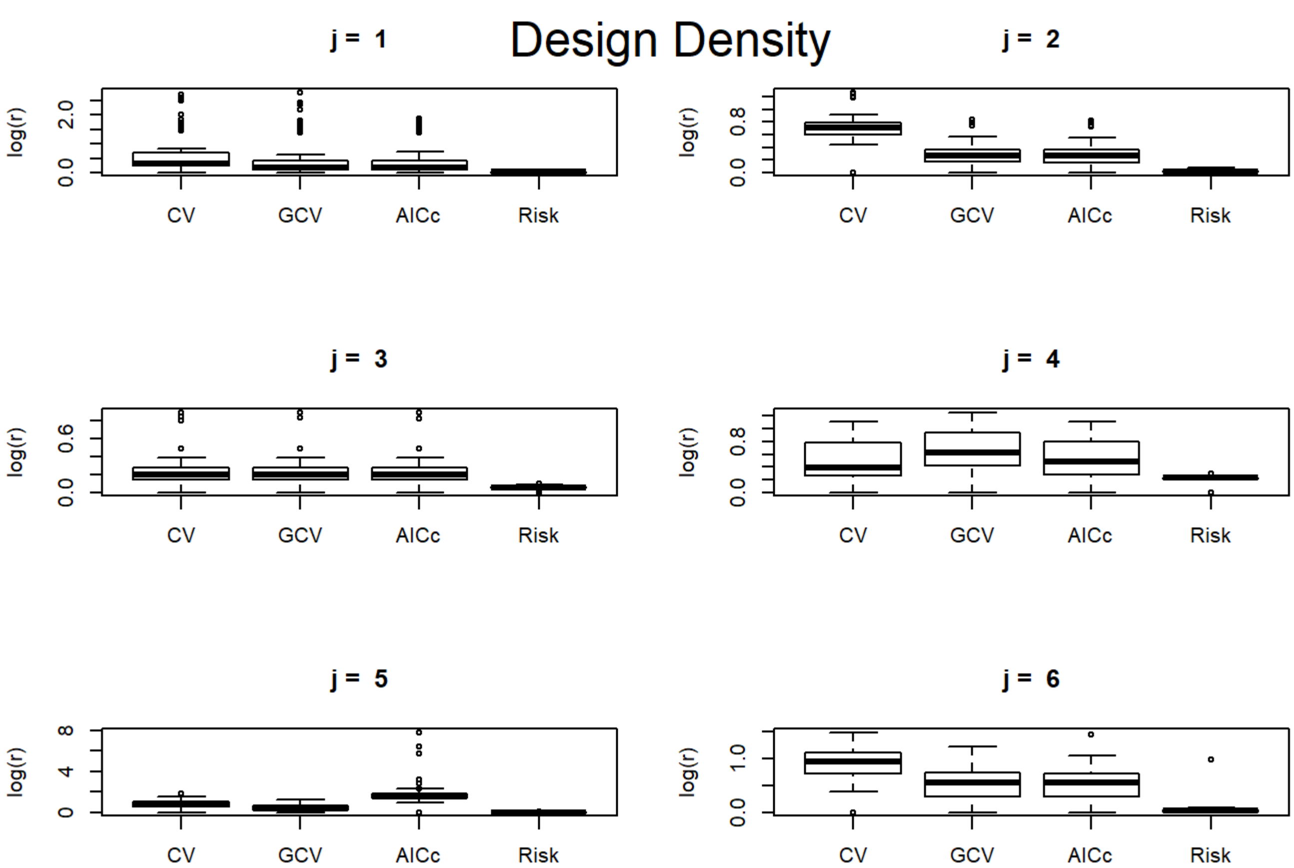
HW5_Maheshwari

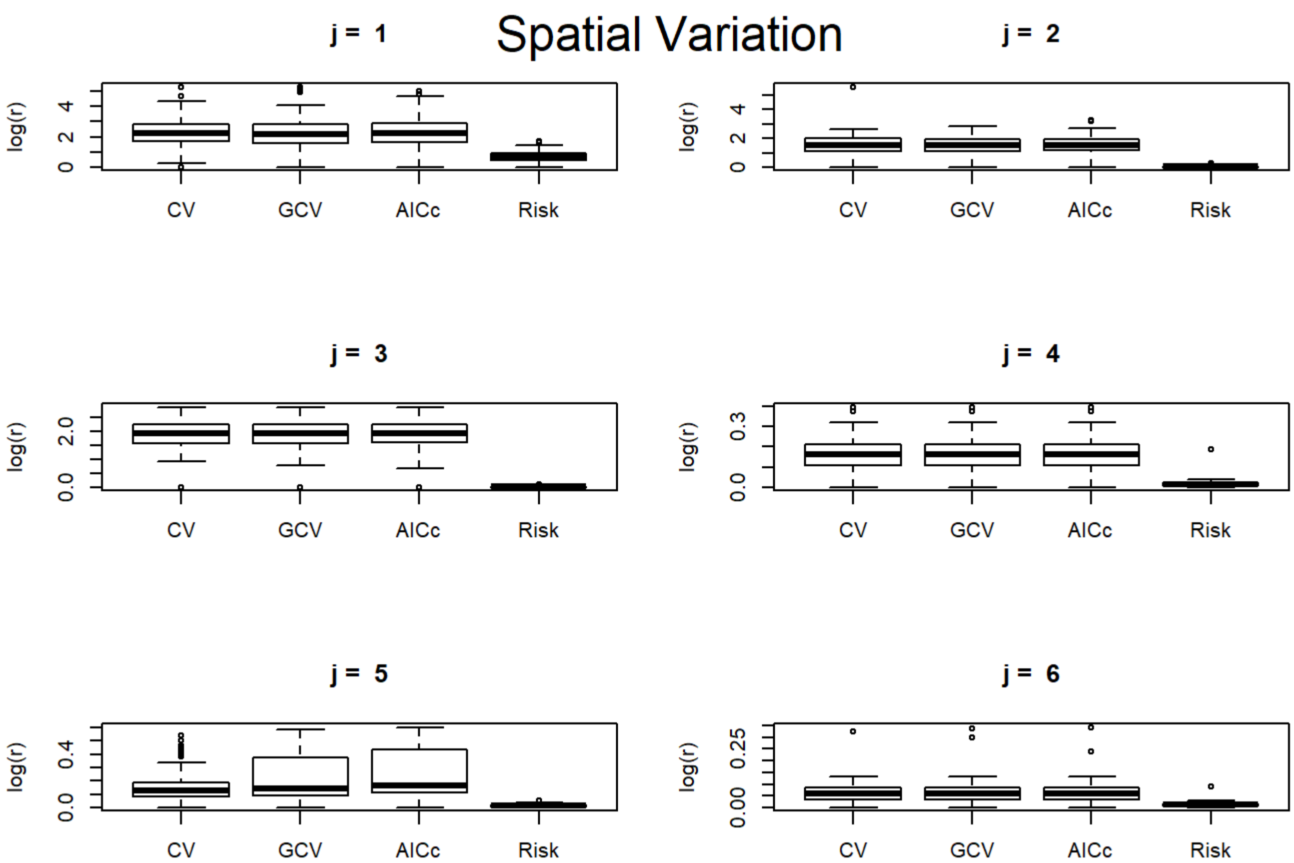
Saurabh Maheshwari - 100%

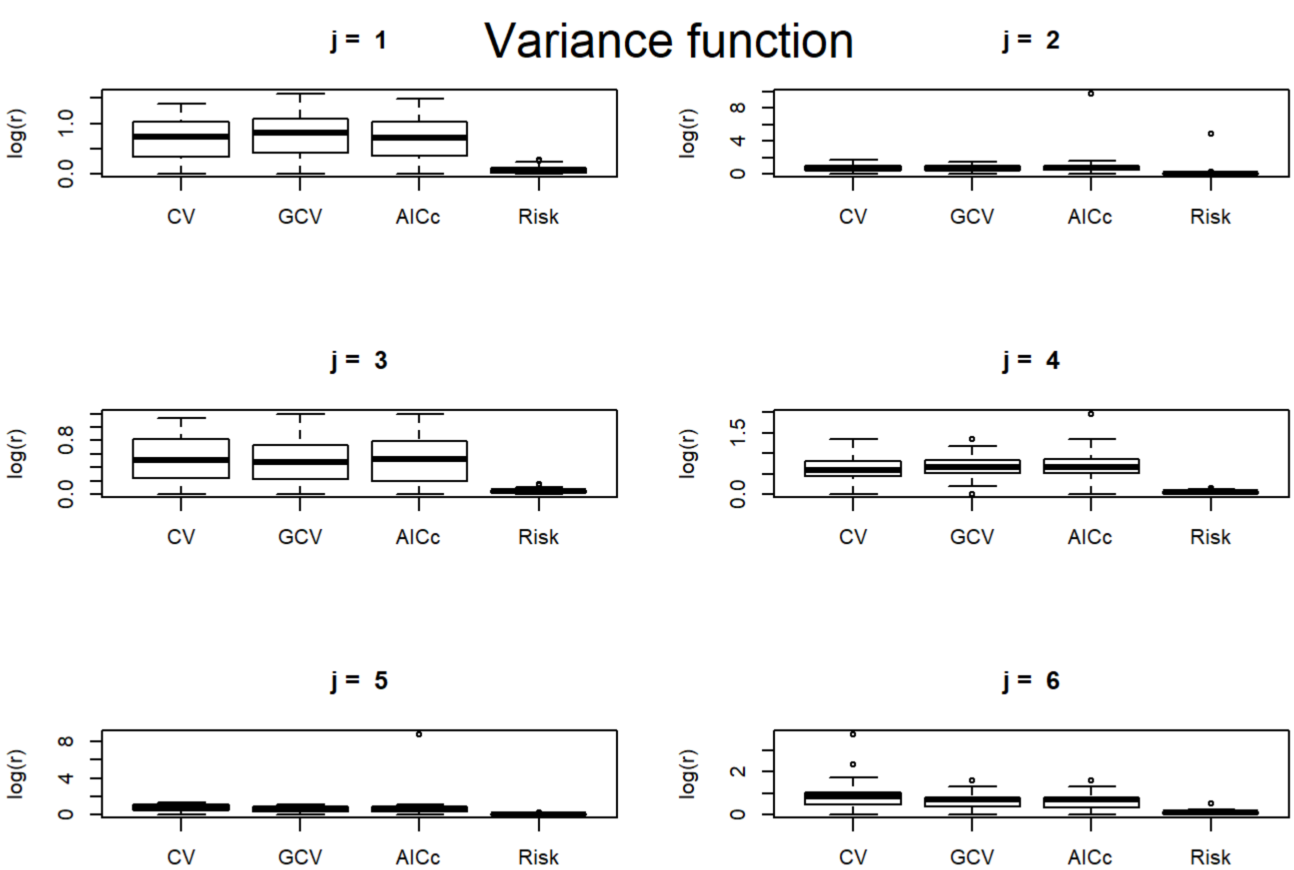
June 1, 2018

Boxplots









The λ chosen using the minimization of L_2 risk shows the most reliable results as the $log_e(r)$ value is the least for this method for almost all the distributions. For CV, GCV and AICc, the $log_e(r)$ values are certainly comparable, the mean value being less than 0.5 for most of the cases makes these approaches reasonable. Among these 3, AICc perform better or equivalent to CV and GenCV for most of the cases. Though, to create boxplots, for each distribution only 80 iterations are considered due to computational limitations, increasing the number of iterations is recommended to get reliable results.

Solution:-

a), b) Penalized Regression splines were fitted using the experimental setup attached in the paper. For all four different setups, namely, Noise level, Design density, spatial variation and Variance function, 6 different conditions were considered. λ for each of these setups was optimized using crossvalidation, AICc and generalized crossvalidation approaches. Finally, a boxplot of $\log_e r$ was created for each distribution, where,

$$r = \frac{\|f - \hat{f}_\lambda\|^2}{\min_\lambda \|f - \hat{f}_\lambda\|^2}$$

c) To prove: $E\|y-\hat{f}_\lambda\|^2 = \|(I-H_\lambda)f\|^2 + \sigma^2\{tr(H_\lambda H_\lambda^T) - 2tr(H_\lambda) + n\}$

L.H.S: $E\|y-\hat{f}_\lambda\|^2 = E\|y - H_\lambda y\|^2 = E\|(f+e) - H_\lambda(f+e)\|^2$

$$= E\|(I-H_\lambda)f + (I-H_\lambda)e\|^2$$

$$= \|(I-H_\lambda)f\|^2 + E(\|(I-H_\lambda)e\|^2) + 2[\underbrace{E((I-H_\lambda)f.(I-H_\lambda)e)}_{0}]$$

$$= \|(I-H_\lambda)f\|^2 + E(e^T(I-H_\lambda)^T(I-H_\lambda)e)$$

$$= \|(I-H_\lambda)f\|^2 + E(tr(e^T(I-H_\lambda)^T(I-H_\lambda)e))$$

$$= \|(I-H_\lambda)f\|^2 + E(tr((I-H_\lambda)^T(I-H_\lambda)ee^T))$$

$$= \|(I-H_\lambda)f\|^2 + \sigma^2.E(tr((I-H_\lambda)^T(I-H_\lambda)))$$

$$= \|(I-H_\lambda)f\|^2 + \sigma^2\{n - 2tr(H_\lambda) + tr(H_\lambda H_\lambda^T)\}$$

To estimate the risk estimator

$$\begin{aligned}
\text{risk}(\lambda) &= E\|f - f_\lambda\|^2 \\
&= E\|f - H_\lambda y\|^2 \\
&= E\|f - H_\lambda f - H_\lambda e\|^2 \\
&= E\|(I - H_\lambda)f - H_\lambda e\|^2 \\
&= \|(I-H_\lambda)f\|^2 + E\|H_\lambda e\|^2 - 2\,\underbrace{E\|(I-H_\lambda)f \cdot H_\lambda e\|}_{0} \\
&= \|(I-H_\lambda)f\|^2 + E\left(\text{tr}\left(e^T H_\lambda^T H_\lambda e\right)\right) \\
&= \|(I-H_\lambda)f\|^2 + \sigma^2\, \text{tr}\left(H_\lambda H_\lambda^T\right)
\end{aligned}$$

Thus,

$$\begin{aligned}
\text{risk}(\lambda) &= E\|y - \hat{f}_\lambda\|^2 - \sigma^2\{n - 2\text{tr}(H_\lambda) + \text{tr}(H_\lambda H_\lambda^T)\} \\
&\quad + \sigma^2\, \text{tr}(H_\lambda H_\lambda^T) \\
&= E\|y - \hat{f}_\lambda\|^2 - \sigma^2\{n - 2\text{tr}(H_\lambda)\}
\end{aligned}$$

Also, $\hat{\sigma} = \sqrt{\dfrac{\Sigma(y-\hat{f})^2}{n-1}}$ is considered as a reasonable estimate for σ.

This approach has been implemented in R, the code has been attached and the boxplots are prepared for comparing the $\log_e r$ values.