Robust Parameter Selection: Understanding Peak Selection in KDE

Jeff Tuche :)

1 Why Not Just Take the Highest Peak?

At first glance, selecting the **highest** peak in the KDE distribution might seem like the best approach. However, the **height of a peak alone does not determine robustness**. Several key factors contribute to the **stability** of a parameter:

1. **Peak Height** - A higher peak suggests many robust parameter values exist around it.
2. **Peak Width (Flatness of the Peak)** - A wider peak means a range of values around the peak are equally good, indicating more stability.
3. **Steepness of the Peak** - A very sharp, isolated peak may indicate a highly **sensitive** parameter where small deviations significantly impact performance.
4. **Multiple Peaks** - If multiple peaks exist, we should prefer the **widest** one over the absolute highest to ensure resilience.

2 How the Selection Works

1. **Step 1:** Apply Kernel Density Estimation (KDE) to the stable parameters and detect peaks.
2. **Step 2:** Identify the **highest peak** and other **local peaks** within a reasonable range.
3. **Step 3:** Among the detected peaks, prefer:
    - The **widest** peak (suggesting more robustness).
    - The **most stable slope** (avoiding steep, sensitive regions).
    - The peak **closest to the median of stable values** (to avoid outliers).

3 Concrete Example

Consider the KDE plot in Figure 1, where two prominent peaks exist:

- **Peak at 0.20**: This is the highest peak, but it is **narrow and steep**, suggesting that minor fluctuations can result in instability.
- **Peak at 0.19**: This peak is slightly lower but has a **wider distribution**, meaning the parameter values around 0.19 are more forgiving and robust.

By selecting **0.19** instead of **0.20**, we ensure that our parameter is **not only high-performing but also resilient to small variations**.

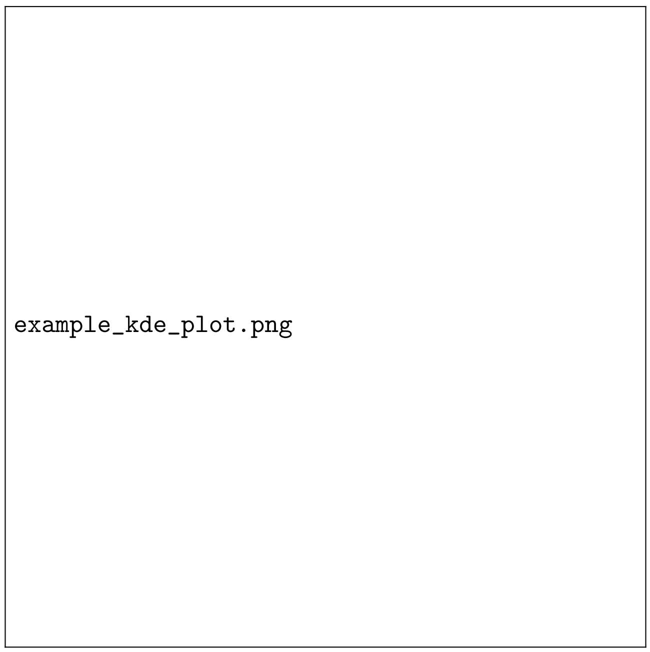

Figure 1: Example KDE plot showing peak selection strategy. The selected value at 0.19 is chosen based on robustness rather than absolute peak height.

4 How to Quantify Stability?

To select the best parameter, we use the following criteria:

1. **Peak Prominence:** If two peaks are close in height, choose the one with a **wider** range.
2. **Peak Width:** Prefer peaks where density remains high across a larger range.

3. **Slope Analysis:** Avoid peaks that are **too steep**, as they indicate high sensitivity.
4. **Median Check:** The final value should **not be an outlier** and should align with the median of stable values.

5 What Can Be Improved?

If we want to adjust the selection strategy, we could:

- **Always pick the highest peak**, ignoring width.
- **Use a weighted function** combining peak height and width.
- **Add a threshold:** If two peaks are close in height, always select the wider one.

6 Conclusion

The current method ensures that parameters are not only high-performing but also robust against fluctuations. However, if a different approach is desired, we can adjust the peak selection criteria accordingly.

Would you like to modify the selection method? Possible adjustments include:

- Selecting strictly the highest peak.
- Weighting peaks based on prominence and width.
- Allowing user preference between peak height vs. width.