The Kalman Filter

By Varun Shah, ECE '21

Introduction

Humans recognize and react to their environment by interpreting a vast array of signals that we receive through our five senses. Each sense tells us something different about the environment around us, and each sense has individual strengths and weaknesses. By combining the information received by each sense, humans can interpret the world around them far better than they could using only one sense at a time.

The ability to combine multi-sensory data has been translated from the natural world to technical applications using a technique known as *sensor fusion*. A system can take in raw data from many sources and combine them to improve the system's understanding of its environment, so the system can predict what will occur next. There are many factors affecting a system that cannot be numerically accounted for, creating *uncertainty* in the received information. The goal of using a sensor fusion algorithm is to combine many pieces of uncertain information to form a prediction that is less uncertain than a prediction made based on a single piece of information.

Kalman Filter

One of the most common sensor fusion algorithms is the Kalman Filter. The algorithm is an efficient and accurate means to estimate the physical parameters, or state, of the system, such as position and velocity. The Kalman Filter equations work to minimize the discrepancies between the true state of the system and the state predicted by the model. Doing so makes the algorithm a reliable method even when running on data that includes noise.

The Kalman Filter is useful in a wide variety of applications because of a few key qualities. First, it is a recursive algorithm, relying only on information from the current state and the previous state (Kalman, 1960). For this reason, it is fast and efficient enough to perform the calculations in real-time, even when using embedded processors that may not be as powerful as a traditional computer.

Under some conditions, the most basic form of the Kalman Filter is the optimal choice of algorithm. First, the system must be *linear*. Without linearity, a more general form of the Kalman Filter can be used, but may not be the best tool for that particular case. Second, the noise that causes the uncertainties must be additive white Gaussian noise (Kim and Bang, 2018). In most circumstances, designers can assume they are dealing with noise of this type, as it is extremely prevalent in real-world situations.

Using a Kalman Filter

The Kalman filter functions via two phases: the prediction phase, and the update phase. Both phases rely on producing *probability distributions*. These are graphs that show the possible range of values for a chosen variable, and their relative likelihoods. In this case, the variable (or variables) being described by the distribution are those that make up the state of the system.

The prediction phase works via a known numerical model of the system. In a system analyzing position and velocity, this model could use basic physics equations. Through this model, a prediction of where the object will be and how fast it is moving is calculated.

The result of the prediction phase then feeds into the update phase. The update phase serves to refine the estimate using measurements. Each measurement comes from a sensor that may have been affected by noise. This uncertainty produces another distribution. This distribution is combined with the distribution from the prediction phase to produce the best possible estimate of the current state of the system.

Prediction Phase

The prediction phase is meant to take into account the known or controllable influences on the system. It intends to combine the previous state of the system with any known changes in the system to produce its prediction of the current state. For an example of a moving car using position and velocity, the prediction phase will produce an estimate based on the last known location and speed of the car, as well as changes made by the driver. The driver could have braked, accelerated, or changed direction, but all of these influences are known control forces. That means that their effects can be modeled via mathematical formulas.

The prediction phase must then account for noise and those influences on the system that are not controllable and unknown. Accounting for these factors introduces uncertainty into the model: the value calculated based on known effects is the most likely outcome, but other outcomes may also be possible if the system is affected by external forces. A tire could slip or there could be wind or weather conditions changing the speed or direction of the system.

The distribution that is produced by this phase is described by the expected value from known control forces, and the variance introduced by unknown influences.

Update Phase

The update phase intends to improve the estimate by taking a sensor reading and comparing it to the estimate made in the previous phase. Similar to the prediction phase, the update phase obtains a value that it assumes to be the most likely, and produces a distribution around that value by taking into account the uncontrollable factors that might be affecting that value. The major difference is that the value is obtained from a sensor reading instead of a mathematical model.

Any sensor reading is inherently imperfect. Sensors are subject to electromagnetic noise, sensitive to changes in temperature and other operating conditions, and have a set range of values for which they can give accurate readings. All of these factors contribute to the uncertainty, and therefore variance, in the distribution produced by the update phase.

Combining Distributions

The two distributions produced represent two probabilities: the probability that our sensor reading is some value, and the probability that our estimate thinks that value is the one we should obtain. The final step is to find the probability that *both* of these two probabilities are true. The distribution this calculation will produce has a much lower uncertainty than either of the original two distributions, giving a best guess of the state considering all the information available. This is shown below. The red distribution is that from the prediction phase and the blue is that from the update phase. The overlap between them, highlighted in pink is the probability distribution that must be calculated.

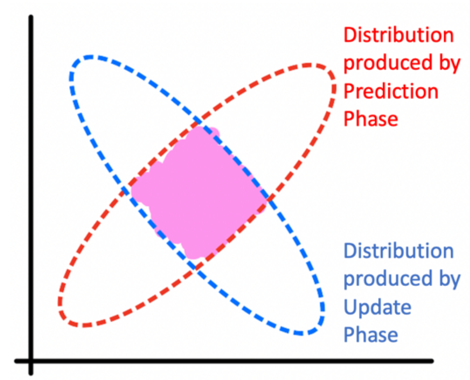

Figure 1. Graphically showing the overlap between the two phases, hand-drawn by the Author (2021)

Just as in basic probability, to obtain the probability distribution of two events both occurring, the distributions of each event are multiplied together. The result is a new distribution, the center of which is a weighted average of the original two distributions' centers, where the more weight is given to the distribution with less noise. The variance in the new distribution is greatly reduced compared to the variance of either of the original two distributions alone. The new distribution is the optimal estimate for the state, and, with the reduction in variance, is an accurate representation of the current state of the system.

To get to the next time step from the current state, i.e., when the algorithm iterates, the optimal estimation of the state is fed back into a new prediction phase. The optimal estimation is taken as fact, and used as the previous state of the system, and the algorithm proceeds as described for the next point in time.

Conclusion

The Kalman Filter is a powerful tool, even in its most basic form. It efficiently uses only the most recent state and the current sensor data to produce a very accurate prediction of the current state of the system. It is applicable across a wide variety of use cases in its base form, and a huge number more in its other forms.

That being said, its use may be limited for the Smart Frisbee Project. One reason is that the motion of a frisbee cannot be modeled with basic physics, as it generates lift during its flight. Another reason is that we may not require such a powerful algorithm to find the information that we hope to present to the user: The position of the disc and its speed during flight is not as important as the instantaneous velocities during or immediately after the throw.

An examination of the Kalman Filter did provide ideas for ways in which uncertainty can be reduced using the wide range of information received by the sensors on the frisbee. A similar process of combining distributions to obtain a more accurate estimate can be used. In this case, however, the distributions come from two or more different sensors that tell us the same information in different forms, instead of an estimate and a sensor reading.

References

1. Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME--Journal of Basic Engineering*, 82, 35-45.

2. Babb, T. (n.d.). How a Kalman Filter Works, in Pictures. Retrieved from https://www.bzarg.com/p/how-a-kalmanfilter-works-in-pictures/

3. Kim, Y and Bang, H. (2018). Introduction to the Kalman Filter and its Applications. Available from: https://www.intechopen.com/books/introduction-and-implementations-of-the-kalman-filter/introduction-to-kalman-filter-and-its-applications DOI: 10.5772/intechopen.80600

4. Welch, G. and Bishop, G. (2006) An Introduction to the Kalman Filter. *Department of Computer Science, University of Carolina at Chapel Hill*. TR 95-041

5. Elmenreich, W. (2002). An Introduction to Sensor Fusion. *Institut f ̈ur Technische Informatik Vienna University of Technology, Austria.* Available from https://www.researchgate.net/publication/267771481_An_Introduction_to_Sensor_Fusion