

Aprendizaje Bayesiano

Bayesian perception: An Introduction. Wei Ji Ma

Chap.7: Time and learning

Introducción

El aprendizaje es un proceso mediante el cual actualizamos lo que creemos sobre el mundo al rededor nuestro. Con cada observación actualizamos estas creencias, que usualmente se vuelven más precisas (si el mundo no está cambiando).

Ya vimos en la materia cómo integrar información mediante Bayes en casos en los que no cambia el estado del mundo mientras estimamos.

En sistemas dinámicos

Integramos información mientras el mundo cambia. La idea es combinar nuestro conocimiento de la dinámica de un proceso con las señales que percibimos. Por ejemplo: combinando el conocimiento de nuestra orientación y velocidad podemos en conjunto con señales intermitentes de un GPS estimar nuestra posición de manera continua.

Los algoritmos bayesianos de filtering alternan dos pasos: aplicación de dinámica y aplicación de la regla de Bayes a medida que llega nueva información.

Hidden Markov Models

En este caso, para modelar procesos en el tiempo, se asume que el estado futuro del mundo depende del estado presente.

En este sentido, las dinámicas dan las reglas que rigen las transiciones entre estados. Más precisamente, dan la distribución de las transiciones a un estado $i+1$ dado el i . De esta forma, combinadas con el estado de las creencias en i , obtendremos el prior para el nuevo estado $i+1$ (antes de la nueva observación).

En el momento de la inferencia, actualizamos las estimaciones para el estado $i+1$ combinando la información provista por una nueva observación y el prior anterior utilizando la regla de Bayes. Finalmente, a este nuevo estado de las creencias se le aplican las dinámicas y el proceso se repite.

Podemos destacar la aplicación de HMMs en modelos para el procesamiento de habla.

Filtros de Kalman

Siguen la misma idea que los HMM pero para estados continuos. El modelo generativo asume una dinámica lineal para gaussianas con ruido:

$$x_{t+1} = Ax_t + \eta$$

Con η una normal centrada en 0 representando el ruido y x una variable gaussiana.

En cada paso, antes de una observación se marginaliza el próximo valor de x sobre todos los valores que pudo haber tenido en el instante anterior. Por tratarse de una convolución de gaussianas el resultado de ajustar a la dinámica del proceso es una gaussiana (vamos a profundizar todas estas cuentas).

Para el segundo paso, como los filtros de Kalman toman priors y likelihoods gaussianas, el resultado de incorporar una nueva observación es también una gaussiana cuya varianza y media se actualiza siguiendo fórmulas conocidas. En el caso de la media se depende de una variable de peso asociada a la observación W (ganancia de Kalman) que decrece con el nivel de conocimiento que ya se tiene.

Smoothing

La idea es usar información futura para mejorar el aprendizaje. La única diferencia en los modelos es que la probabilidad de los estados del mundo no es condicional solo en el pasado y presente sino que también del futuro. La propiedad de Markov simplifica en gran medida estas dependencias (cuanto profundizamos de esto depende de cuánto ocupe lo demás).

Propuestas para implementación

Posibilidades:

- Replicar alguno de los ejemplos del capítulo pero con un modelo distinto
- Implementar alguno de los modelos de la sección de problemas del capítulo. Uno en particular propone un set de datos dedicado a científicos que busquen modelar y analizar movimientos de mano/brazo utilizando filtros de Kalman. En función de lo que discutimos el Lunes, la intención es armar una resolución guiada del ejercicio en R.