Filtrelemek

Filtreler dis dunyadaki bir aksiyon hakkinda elde edilen gurultulu sinyalleri, tersine cevirerek arka plandaki aksiyon hakkinda hesaplama yapabilmemizi saglar. Mesela Kalman Filtreleri (KF) icin gizlenmis konum bir robotun nerede oldugu, bir senetin fiyati gibi bir sey olabilir, gizli konum bilgisi x_t degiskeninde o konum hakkindaki gurultulu olcum y_t icindedir. Hem gizli konumlar arasindaki gecis, hem de olcumun gurultusu lineer bir fonksiyon uzerindendir.

$x_{t+1} = Ax_t + v$

$y_t = Hx_t + w$

v ve w'in dagilimi Gaussian'dir ve kovaryans sirasiyla Q ve R icindedir.

Zaman faktorunu de dahil etmek gerekirse;

$\hat{x}_t^t = E[x_t|y_0, .., y_t]$

$P_t^t = E[(x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})'|y_0, ..., y_t]$

Filtremenin amaci x_{t+1} ve P_{t+1} hesabini yeni bir olcum y_{t+1} uzerinden yapmak olacak. "Gizli" x_t derken bunu kastediyorduk, bu deger bize verilmiyor, sadece xt ve x_{t+1} arasindaki gecisin nasil oldugunu biliyoruz, gurultunun nasil eklendigini biliyoruz, ama bunlarin bilsek bile elde bir suru bilinmeyen var. Filtrelemenin matematiksel numaralari sayesinde bunu hesaplayabiliyor olacagiz. Yani yapmamiz gereken "oku tersine cevirmek", yani x_t'nin y_t uzerindeki sartsal bagliligini (conditional dependence) ortaya cikartmak, bunu y_t'nin x_t'ye olan sartsal bagimliligini tersine cevirerek yapmak. Ana denklemin iki tarafinin da beklentisini (expectation) alalim:

$E\ x_{t+1} = \hat{x}_{t+1} = A\mu_t = A\hat{x}_t$

Simdi iki tarafin kovaryansini alalim ve P_t'yi $cov\ x(t)$ olarak belirtelim:

$P_{t+1} = AP_tA' + Q$

Bu gecis "zaman guncellemesi" olarak adlandirilir. Normal dagilimlari t anindan $t+1$ anina gecirmemizi saglar. y iceren formullerde benzer bir durum var.

$\hat{x}_{t+1}^t = Ax_t^t$

$P^t_{t+1} = AP^t_t A' + Q$

$y_{t+1} = Cx_{t+1} + w_t$

$E[y_{t+1}|y_0, .., y_t] = E[Cx_{t+1} + w_t|y_0, .., y_t]$

$\hat{y}^t_{t+1} = C\hat{x}_{t+1}$

Kovaryans icin benzer durum

$E[(y_{t+1} - \hat{y}^t_{t+1})(y_{t+1} - \hat{y}^t_{t+1})'|y_0, ..., y_t] = C^t_{t+1}C' + R$

Simdi daha zor is olan oku tersini cevirmeye gelelim. Eger amacimiz p(xt —yt) denklemini elde etmek ise o zaman bu iki degiskeni iceren birlesik dagilimi (joint distribution) elde etmek zorundayiz. Iki Gaussian'in birlesiminin yeni bir Gaussian oldugunu biliyoruz, o zaman hem x_t hem de y_t'in kendisi cok boyutlu birer Gaussian olduklari icin onlarin birlesimi $p(x_t|y_t)$'in hakikaten devasa bir Gaussian olacagini tahmin edebiliriz.

x_t ve y_t'in birlesimi olan Gaussian'i bulmak demek, bu Gaussian'in ortalamasini (mean) ve kovaryansini bulmak demektir cunku bir Gaussian ortalama ve kovaryansi ile net bir sekilde tanimlanabilir bir seydir. Bir numara yapalim, ve $y_t = Cx_t + w_t$'yi $z = Hu$ seklinde yazalim. Sonra

$\left[\begin{array}{c} x_t \\ y_t \end{array}\right], H = \left[\begin{array}{cc} I & 0 \\ C & I \end{array}\right], u = \left[\begin{array}{c} x_t \\ w_t \end{array}\right]$

Boylece daha basit bir denklemin kovaryansini alabiliriz

$cov(z) = H\ cos(u)H'$

$cov(u) = \left[\begin{array}{cc} P_t & 0 \\ 0 & R \end{array}\right]$

Tam carpim suna esit

$\left[\begin{array}{cc} I & 0 \\ C & I \end{array}\right] \left[\begin{array}{cc} P_t & 0 \\ 0 & R \end{array}\right] \left[\begin{array}{cc} I & C' \\ 0 & I \end{array}\right]$

bunun sonucu ise

$\left[\begin{array}{cc} P_t & P_tC' \\ CP_t & CP_tC' + R \end{array}\right]$

Bunu baglantisal denklem icin ve ortalamayi icerecek sekilde yazabiliriz

$$\left[\begin{array}{c} \hat{x}_t^t \\ C\hat{x}_t^t \end{array}\right], \left[\begin{array}{cc} P_t^t & P_t^t C' \\ CP_t^t & CP_t^t + R \end{array}\right]$$

Ayni sekilde x_{t+1}, y_{t+1} birlesik dagilim icin

$$\left[\begin{array}{c} \hat{x}_{t+1}^t \\ C\hat{x}_{t+1}^t \end{array}\right], \left[\begin{array}{cc} P_{t+1}^t & P_{t+1}^t C' \\ CP_{t+1}^t & CP_{t+1}^t C' + R \end{array}\right]$$

Simdi x_{t+1}^{t+1} 'in ortalama ve varyansi icin parcali Gaussian kavramini anlatmaliyiz. Bir n boyutlu Gaussian daha kucuk boyutlardaki p ve q alt Gaussian'lara parcalanabilir (tabii ki $n = p + q$). Yani su ifade kullanilabilir

$$\mu = \left[\begin{array}{c} \mu_1 \\ \mu_2 \end{array}\right], \Sigma = \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array}\right]$$