

# Vehículos Espaciales y Misiles

## Tema 3: Control y determinación de la actitud Parte I: Estimación de actitud

Rafael Vázquez Valenzuela

Departamento de Ingeniería Aeroespacial  
Escuela Superior de Ingenieros, Universidad de Sevilla [rvazquez1@us.es](mailto:rvazquez1@us.es)

9 de abril de 2014



## Estimación: sensores

- La mayor parte de los sensores de actitud se pueden clasificar, a grandes rasgos, en dos tipos:
  - Sensores que determinan un vector  $\vec{v}$  en los ejes cuerpo, es decir,  $\vec{v}^B$  (en realidad lo determinan en los “ejes sensor”, pero la transformación a ejes cuerpo es conocida). Además dicho vector se supone conocido en los ejes de referencia (inerciales u órbita), es decir,  $\vec{v}^N$  es conocido. Es necesario tener dos o más medidas simultáneas para poder resolver el problema sólo con este tipo de sensores.
  - Sensores que miden la velocidad angular  $\vec{\omega}^B$  con respecto al sistema de referencia inercial. Se recupera la actitud usando esta medida para integrar las ecuaciones diferenciales cinemáticas, pero los errores se acumulan generando una cierta deriva: es necesario complementar con alguna medida del tipo anterior.
- Todos los sensores además tendrán error de medida; este error se puede modelar como un sesgo (error constante) más un ruido blanco (aleatorio, con una cierta covarianza).



## Descripción estadística del error

- Consideremos por ejemplo el caso del error de medida de un giróscopo:  $\vec{\omega} = \hat{\vec{\omega}} + \delta\vec{\omega}$ , donde  $\delta\vec{\omega}$  son errores de medida y  $\hat{\vec{\omega}}$  la medida obtenida.
- Una componente de  $\delta\vec{\omega}$ , por ejemplo  $\delta\omega_1$ , puede tener el siguiente aspecto:



- Es imposible conocer el valor con exactitud.
- Se observa que cambia con el tiempo.
- Por tanto, se representan sus propiedades usando la estadística.
- La teoría de procesos estocásticos nos permitirá analizar como el ruido de los sensores se propaga como error de estimación.



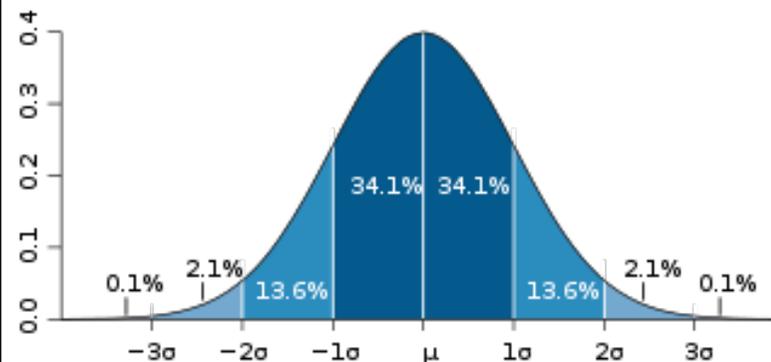
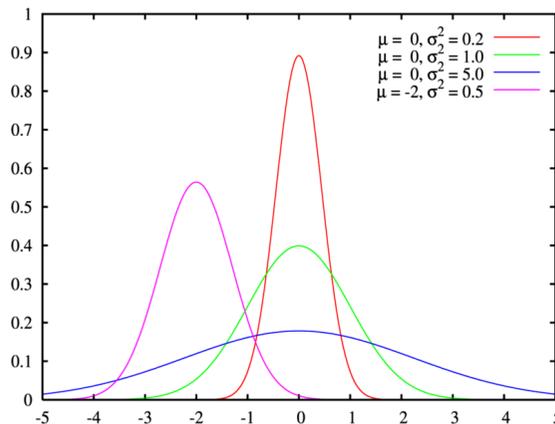
## Variables aleatorias continuas unidimensionales

- Sea una variable aleatoria  $X \in \mathbb{R}$  continua.
- Recordemos que la función de distribución  $F(x)$  es la probabilidad de que  $X \leq x$ , que se escribe como  $F(x) = P(X \leq x)$ .
- La función de distribución se calcula mediante la función de densidad  $f(x)$ :  $F(x) = \int_{-\infty}^x f(y)dy$ .
- Se define el operador esperanza matemática actuando sobre la función  $g(x)$  como  $E[g(X)] = \int_{-\infty}^{\infty} g(y)f(y)dy$ . Se trata de un operador lineal, de forma que  $E[\alpha_1 g_1(X) + \alpha_2 g_2(X)] = \alpha_1 E[g_1(X)] + \alpha_2 E[g_2(X)]$ . Los dos casos importantes son:
  - Media:  $m(X) = E[X] = \int_{-\infty}^{\infty} yf(y)dy$ .
  - Varianza:  $V(X) = E[(X - m(X))^2] = E[X^2] - (E[X])^2$ .
  - Desviación típica  $\sigma$ , la raíz cuadrada de la varianza,  $\sigma = \sqrt{V(X)}$ .



# Distribución normal o gaussiana I

- Es la distribución más usada en estadística. Se escribe  $X \sim N(m, \sigma^2)$  y su función de densidad es 
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{Exp} \left( -\frac{(x-m)^2}{2\sigma^2} \right).$$
- Intervalos de confianza: si  $X \sim N(m, \sigma^2)$ :
  - Intervalo 1- $\sigma$ :  $P(X \in [m - \sigma, m + \sigma]) = 68,3 \%$ .
  - Intervalo 2- $\sigma$ :  $P(X \in [m - 2\sigma, m + 2\sigma]) = 95,45 \%$ .
  - Intervalo 3- $\sigma$ :  $P(X \in [m - 3\sigma, m + 3\sigma]) = 99,74 \%$ .



## Distribución normal o gaussiana II

- El **teorema central del límite** dice que la suma de variables aleatorias (con cualquier tipo de distribución) tiende en media a la normal. Puesto que los errores a gran escala provienen de la suma y acumulación de muchos errores a pequeña escala, esto justifica el uso de la normal como modelo para errores.
- Una propiedad importante de la normal es que la suma de normales es de nuevo normal, es decir, si  $X \sim N(m_x, \sigma_x^2)$  e  $Y \sim N(m_y, \sigma_y^2)$  y son independientes, entonces si  $Z = X + Y$  se tiene que  $Z \sim N(m_x + m_y, \sigma_x^2 + \sigma_y^2)$ .
- Por tanto  $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$ , es decir, la desviación típica de la suma de errores es **la raíz cuadrada de la suma de los cuadrados de las desviaciones típicas de los errores**.
- Esta regla, conocida como Root-Sum-of-Squares (RSS) es muy importante.



## Variables aleatorias continuas multidimensionales

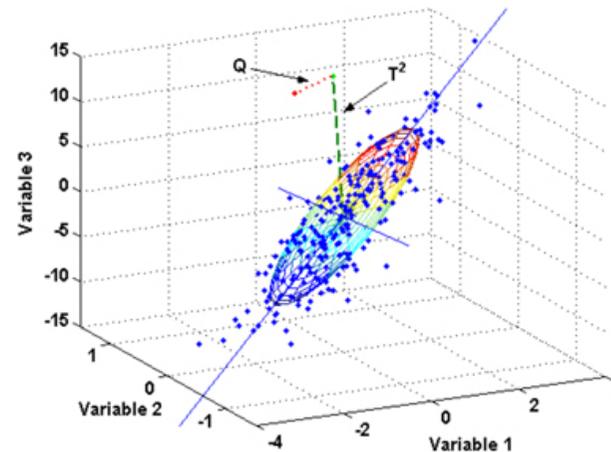
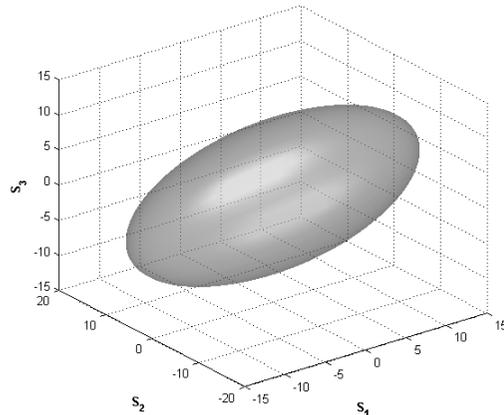
- Sea una variable aleatoria  $\vec{X} \in \mathbb{R}^n$  continua multidimensional.
- Cada componente de  $\vec{X}$  sigue una distribución unidimensional.
- Como en el caso unidimensional, se define una función de distribución conjunta, que se calcula mediante la función de densidad  $f(\vec{x})$ .
- Igualmente  $E[g(\vec{X})] = \int_{\mathbb{R}^n} g(\vec{y})f(\vec{y})d\vec{y}$ . Los dos casos importantes son:
  - Media:  $\vec{m}(\vec{X}) = E[\vec{X}] = \int_{\mathbb{R}^n} \vec{y}f(\vec{y})d\vec{y}$ .
  - Covarianza:  $Cov(\vec{X}) = E[(\vec{X} - m(\vec{X}))(\vec{X} - m(\vec{X}))^T] = \Sigma$ . Es una matriz simétrica y definida positiva. Los valores de la diagonal representan la varianza de cada componente de  $\vec{X}$ , y los otros valores la correlación entre dos componentes de  $\vec{X}$ . Se tiene  $\Sigma = E[(\vec{X}\vec{X}^T] - m(\vec{X})m(\vec{X})^T$ .
- Por ejemplo, para  $n = 3$  y escribiendo  $\vec{X} = [X, Y, Z]$ :

$$\Sigma = \begin{bmatrix} \sigma_x^2 & E[(X - m_x)(Y - m_y)] & E[(X - m_x)(Z - m_z)] \\ E[(X - m_x)(Y - m_y)] & \sigma_y^2 & E[(Y - m_y)(Z - m_z)] \\ E[(X - m_x)(Z - m_z)] & E[(Y - m_y)(Z - m_z)] & \sigma_z^2 \end{bmatrix}$$



## Distribución normal multivariante I

- Se escribe  $\vec{X} \sim N_n(\vec{m}, \Sigma)$  y su función de densidad es  $f(\vec{x}) = \frac{1}{\text{Det}(\Sigma)(2\pi)^{n/2}} \text{Exp} \left( -\frac{1}{2}(\vec{x} - \vec{m})^T \Sigma^{-1}(\vec{x} - \vec{m}) \right)$ .
- Los intervalos de confianza son ahora regiones de  $\mathbb{R}^n$ , definidos por  $P(\vec{X} \in \Omega) = P_\Omega$ .
- La forma de estas regiones de confianza es de elipsoides, descritos por la ecuación  $(\vec{x} - \vec{m})^T \Sigma^{-1}(\vec{x} - \vec{m}) = d^2$ , donde  $d$  depende de  $P_\Omega$ . Cuanto mayores sean los valores de los autovalores de  $\Sigma$ , mayor será el elipsoide. Las direcciones de los ejes del elipsoide vendrán dados por los autovectores de  $\Sigma$ .



## Distribución normal multivariante II

- Tomando un ejemplo aeronáutico clásico en navegación aérea, si por ejemplo describimos el error en posición de un avión en ejes cuerpo,  $\delta \vec{r}^b = [\delta x \ \delta y \ \delta z]^T$ , como una normal multivariante con  $n = 3$ , de media cero (centrada en el avión) y con matriz de covarianzas:

$$\Sigma = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix}$$

- Entonces podemos asimilar el movimiento del avión al movimiento del elipsoide, que representa una región de incertidumbre donde se puede encontrar el avión con gran probabilidad.
- Se verifica que si  $\vec{X} \sim N_n(\vec{m}_x, \Sigma_x)$  e  $\vec{Y} \sim N_n(\vec{m}_y, \Sigma_y)$  y son independientes, entonces si  $\vec{Z} = \vec{X} + \vec{Y}$  resulta  $\vec{Z} \sim N_n(\vec{m}_x + \vec{m}_y, \Sigma_x + \Sigma_y)$ .
- Igualmente  $A\vec{X} + \vec{b}$  donde  $A$  y  $b$  son no-aleatorios verifica que  $A\vec{X} + \vec{b} \sim N_n(A\vec{m}_x + \vec{b}, A\Sigma_x A^T)$ .



## Procesos estocásticos.

- Un proceso estocástico o variable estocástica no es sino una variable aleatoria  $\vec{X}(t)$  que cambia con el tiempo. Los errores de navegación serán este tipo de variables.
- Por tanto la media y la covarianza también varían con el tiempo:  $\vec{m}(t)$ ,  $\Sigma(t)$ .
- Para un proceso, se define la autocorrelación como  $R(t, \tau) = E[\vec{X}(t)\vec{X}(\tau)^T]$ . La autocorrelación permite conocer hasta que punto la historia pasada de  $\vec{X}$  influye en su valor actual.
- **Proceso gaussiano:** Un proceso gaussiano verifica  $\vec{X}(t) \sim N_n(\vec{m}(t), \Sigma(t))$ , es decir, se distribuye como una normal multivariante cuya media y covarianza varían con el tiempo.



## Ruido blanco.

- **Ruido blanco:** Se define como ruido blanco un proceso  $\vec{v}(t)$  que verifica:
  - $E[\vec{v}(t)] = \vec{0}$ .
  - $E[\vec{v}(t)\vec{v}(t)^T] = \sigma^2\text{Id}$ .
  - $R(t, \tau) = E[\vec{v}(t)\vec{v}(\tau)^T] = \delta(t - \tau)\sigma^2\text{Id}$ , donde  $\delta(x)$  vale 1 si  $x = 0$  y 0 en cualquier otro caso.
- La última condición quiere decir que el valor del ruido blanco en un instante es independiente de su valor en cualquier instante anterior.
- **Ruido blanco gaussiano:** Es un proceso que cumple las condiciones anteriores, y además es gaussiano.
- Un buen modelo para las fuentes de error de sensores es  $\delta\vec{e}(t) = \vec{b} + D\vec{v}$ , donde  $\vec{v}$  es ruido blanco gaussiano. El valor de  $\vec{b}$  será la media del error (sesgo, llamado bias en inglés).



## Propagación del error. Caso continuo.

- Consideremos una ecuación diferencial del tipo

$$\dot{\vec{x}} = A\vec{x} + D\vec{b},$$

donde  $\vec{b}$  es ruido blanco gaussiano de covarianza  $\sigma^2 \text{Id}$ , y la condición inicial es también gaussiana, es decir,  $\vec{x}_0 \sim N_n(\vec{m}_0(t), P_0(t))$ . Entonces se tiene que  $\vec{x}$  es un proceso gaussiano,  $\vec{x} \sim N_n(\vec{m}(t), P(t))$ , con media y covarianza evolucionando de la siguiente forma:

$$\begin{aligned}\dot{\vec{m}} &= A\vec{m}, \\ \dot{P} &= AP + PA^T + \sigma^2 DD^T, \\ \vec{m}(0) &= \vec{m}_0, \\ P(0) &= P_0\end{aligned}$$



## Propagación del error. Caso discreto.

- Consideremos una ecuación discreta del tipo

$$\vec{x}_{k+1} = A\vec{x}_k + D\vec{b}_k,$$

donde  $\vec{b}_k$  es ruido blanco gaussiano de covarianza  $\sigma^2 \text{Id}$ , y la condición inicial es también gaussiana, es decir,  $\vec{x}_0 \sim N_n(\vec{m}_0(t), P_0(t))$ . Entonces se tiene que  $\vec{x}_k$  es un proceso gaussiano,  $\vec{x}_k \sim N_n(\vec{m}_k(t), P_k(t))$ , con media y covarianza evolucionando de la siguiente forma:

$$\begin{aligned}\vec{m}_{k+1} &= A\vec{m}_k, \\ P_{k+1} &= AP_kA^T + \sigma^2 DD^T,\end{aligned}$$



## Estimación a partir de observaciones

- En primer lugar consideramos el caso de que tenemos  $n$  (2 o más) sensores que determinan un vector  $\vec{v}_i$ ,  $i = 1, \dots, n$ , en los ejes cuerpo, es decir,  $\vec{v}_i^B$ , que es conocido en los ejes de referencia (inerciales u órbita), es decir,  $\vec{v}_i^N$  es conocido. Puesto que sólo importan las direcciones de los vectores, los consideramos unitarios.
- Por tanto tenemos  $n$  ecuaciones de la forma  $\vec{v}_i^B = C_N^B \vec{v}_i^N$  y tenemos que “despejar” la matriz  $C_N^B$ .
- Para simplificar notación escribamos  $\vec{W}_i = \vec{v}_i^B$ ,  $\vec{V}_i = \vec{v}_i^N$ ,  $A = C_N^B$ . Por tanto tenemos  $n$  ecuaciones  $\vec{W}_i = A \vec{V}_i$  y es necesario resolver  $A$ .
- En el caso de tener 2 vectores, se puede usar un método sencillo llamado método TRIAD. También veremos un método más general que sirve para  $n \geq 2$  llamado Q.



## Método TRIAD

- Partimos de las dos ecuaciones  $\vec{W}_1 = A\vec{V}_1$  y  $\vec{W}_2 = A\vec{V}_2$
- Definamos los siguientes vectores:  $\vec{r}_1 = \vec{V}_1$ ,  $\vec{r}_2 = \frac{\vec{V}_1 \times \vec{V}_2}{|\vec{V}_1 \times \vec{V}_2|}$ , y  $\vec{r}_3 = \frac{\vec{V}_1 \times \vec{r}_2}{|\vec{V}_1 \times \vec{r}_2|}$ . Similarmente:  $\vec{s}_1 = \vec{W}_1$ ,  $\vec{s}_2 = \frac{\vec{W}_1 \times \vec{W}_2}{|\vec{W}_1 \times \vec{W}_2|}$ , y  $\vec{s}_3 = \frac{\vec{W}_1 \times \vec{s}_2}{|\vec{W}_1 \times \vec{s}_2|}$ . Está claro que se cumplirá, por construcción:  $\vec{s}_1 = A\vec{r}_1$ ,  $\vec{s}_2 = A\vec{r}_2$ , y  $\vec{s}_3 = A\vec{r}_3$ .
- Construimos ahora las matrices  $M_{ref} = [\vec{r}_1 \ \vec{r}_2 \ \vec{r}_3]$  y  $M_{obs} = [\vec{s}_1 \ \vec{s}_2 \ \vec{s}_3]$ . Se cumple  $M_{obs} = AM_{ref}$ . Además las columnas de la matriz  $M_{ref}$  son, por construcción, ortornormales, luego  $M_{ref}$  es invertible y además ortogonal. Por tanto:  $A = M_{obs}M_{ref}^T$ .
- Obsérvese que el método no es simétrico, sino que da más “peso” a la medida número 1. Por tanto habrá que usar como primera medida la más “fiable”. En la práctica no se obtendrá la matriz exacta por los errores en los sensores.



## Método Q

- Ahora tenemos  $n$  medidas cumpliendo  $\vec{W}_i = A\vec{V}_i$ . Planteamos el problema de forma similar a un problema de mínimos cuadrados.
- Formulamos la función  $L(A) = \frac{1}{2} \sum_{i=1}^n a_i |\vec{W}_i - A\vec{V}_i|^2$ , donde  $a_i$  son los pesos que se le da a cada medida (cumpliendo  $\sum_{i=1}^n a_i = 1$ ) y planteamos el objetivo matemático de encontrar  $A$  (ortogonal) tal que  $L(A)$  sea mínimo.
- Puesto que, operando:

$$|\vec{W}_i - A\vec{V}_i|^2 = (\vec{W}_i - A\vec{V}_i)^T (\vec{W}_i - A\vec{V}_i) = 2 - 2\vec{W}_i^T A\vec{V}_i,$$

tenemos que:

$$L(A) = 1 - \sum_{i=1}^n a_i \vec{W}_i^T A\vec{V}_i = 1 - g(A),$$

donde  $g(A) = \sum_{i=1}^n a_i \vec{W}_i^T A\vec{V}_i$ . Minimizar  $L(A)$  equivale a maximizar  $g(A)$ .



## Método Q

- Recordando que la matriz  $A$  se puede escribir en función de los cuaterniones usando la ecuación de Euler-Rodrigues, se puede demostrar que  $g(A)$ , escrito en función de cuaterniones, es:

$$g(q) = q^T K q$$

donde la matriz  $K$  se escribe a partir de

$$\begin{aligned} B &= \sum_{i=1}^n a_i \vec{W}_i \vec{V}_i^T, \\ \sigma &= \text{Tr}(B), \\ S &= B + B^T, \\ \vec{z}^\times &= B - B^T \end{aligned}$$

siendo  $K$  una matriz  $4 \times 4$  igual a:

$$K = \begin{bmatrix} \sigma & \vec{z}^T \\ \vec{z} & S - \sigma \text{Id} \end{bmatrix}$$



## Método Q

- Por tanto el problema se reduce a encontrar  $q$  (cuaternión válido de actitud, es decir, de norma 1) tal que  $g(q) = q^T K q$  sea máximo.
- Problema de maximización con restricciones ( $q^T q = 1$ ). Usamos multiplicadores de Lagrange, escribiendo:

$$H = q^T K q - \lambda(q^T q - 1)$$

- Tomando derivada respecto a  $q$  e igualando a cero:

$$\frac{\partial H}{\partial q} = 2q^T K - 2\lambda q^T = 0 \quad \longrightarrow \quad Kq = \lambda q,$$

- Por tanto  $\lambda$  debe ser un autovalor de  $K$  y  $q$  el autovector asociado (de módulo 1). ¿Cuál? Sustituyendo en  $g(q)$ :

$$g(q) = q^T K q = q^T \lambda q = \lambda$$

- Por tanto, el cuaternión solución del problema será el autovector (de módulo 1) asociado al mayor autovalor!



## Matrices de covarianza de los métodos Q y TRIAD

- Para el método TRIAD, la matriz  $C_N^B$  obtenida contendrá errores. La forma de modelar estos errores es usando un vector "de ángulos pequeños". Denotando  $\hat{C}_N^B$  a la matriz estimada, realmente  $\hat{C}_N^B = C_n^{\hat{b}}$ , donde se tiene que:

$$N \longrightarrow \hat{B} \xrightarrow[x^b]{\delta\phi_x} S_1 \xrightarrow[y^{S_1}]{\delta\phi_y} S_2 \xrightarrow[z^{S_2}]{\delta\phi_z} B$$

- Se tiene entonces  $C_N^B = C_{\hat{B}}^B C_N^{\hat{B}}$ , por analogía con las definiciones anteriores definimos  $\delta C_N^B = C_N^B - \hat{C}_N^B = C_{\hat{B}}^B \hat{C}_N^B - \hat{C}_N^B = (C_{\hat{B}}^B - \text{Id}) \hat{C}_N^B$ .
- Suponiendo que los errores  $\delta\vec{\phi} = [\delta\phi_x \ \delta\phi_y \ \delta\phi_z]^T$  son pequeños, se vio que  $C_{\hat{B}}^B = \text{Id} - \delta\vec{\phi}^\times$ .
- Por tanto, la relación entre la "matriz de error"  $\delta C_N^B$  y  $\delta\vec{\phi}$  es  $\delta C_N^B = (\text{Id} - \delta\vec{\phi}^\times - \text{Id}) \hat{C}_N^B = -\delta\vec{\phi}^\times \hat{C}_N^B$ . Y se tiene  $C_N^B = (\text{Id} - \delta\vec{\phi}^\times) \hat{C}_N^B$ .



## Matrices de covarianza de los métodos Q y TRIAD

- Para el método TRIAD, modelamos el error desconocido  $\delta\vec{\phi}$  como una normal multivariante con media cero y covarianza  $P_{\phi\phi}$ . Se demuestra la siguiente relación:

$$P_{\phi\phi} = \sigma_1^2 \text{Id} + \frac{1}{|\vec{w}_1 \times \vec{w}_2|^2} \left( (\sigma_2^2 - \sigma_1^2) w_1 w_1^T + \sigma_1^2 (w_1^T w_2) (w_1 w_2^T + w_2 w_1^T) \right)$$

donde  $\sigma_1$  representa el error (desviación típica) de la primera medida y  $\sigma_2$  el error de la segunda medida. Se observa la mayor influencia del error de la primera variable.

- En el caso del algoritmo Q, puesto que se calcula un error, el error vendrá dado por un cuaternión de error  $\delta q$ , de forma que  $\hat{q} = q \star \delta q$ , tal como se definió en el tema 3:

$$\delta q(\vec{a}) = \frac{1}{\sqrt{4 + \|\vec{a}\|^2}} \begin{bmatrix} 2 \\ \vec{a} \end{bmatrix}$$



## Matrices de covarianza de los métodos Q y TRIAD

- Por tanto es ahora  $\vec{a}$  quien representa el error de actitud (a través del cuaternión de error) y por tanto  $\vec{a}$  sería una normal multivariante de media 0 y covarianza  $P_a$ .
- En el algoritmo Q cada medida tendrá su error representado por su varianza  $\sigma_i^2$ . El error final dependerá de los pesos elegidos en el algoritmo y en general, se demuestra que

$$P_a = \frac{1}{4} \left[ \text{Id} - \sum_{i=1}^n a_i \vec{W}_i \vec{W}_i^T \right]^{-1} \left[ \sum_{i=1}^n a_i^2 \sigma_i^2 \left[ \text{Id} - \vec{W}_i \vec{W}_i^T \right] \right] \left[ \text{Id} - \sum_{i=1}^n a_i \vec{W}_i \vec{W}_i^T \right]^{-1}$$

- Una elección de  $a_i$  sería hacerla proporcional al inverso de la varianza  $\sigma_i^2$ . Como los  $a_i$  tienen que sumar 1, se podría elegir

$$a_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^n \frac{1}{\sigma_j^2}}$$

llegando a

$$P_a = \frac{1}{4} \frac{1}{\sum_{j=1}^n \frac{1}{\sigma_j^2}} \left[ \text{Id} - \sum_{i=1}^n a_i \vec{W}_i \vec{W}_i^T \right]^{-1}$$

- Nota: Si  $\left[ \text{Id} - \sum_{i=1}^n a_i \vec{W}_i \vec{W}_i^T \right]$  no es invertible el problema no tiene solución única por el algoritmo Q.



## Ejemplo 1-D: propagación del error de un giróscopo

- Cuando uno tiene medidas de giróscopos, es necesario integrar las ecuaciones diferenciales cinemáticas.
- Para entender el filtro de Kalman a nivel conceptual, vamos a ejemplificarlo en el caso más sencillo: un único grado de libertad de giro. Por tanto hay un sólo ángulo  $\theta$ , cuya ecuación diferencial cinemática es

$$\dot{\theta} = \omega$$

- Un giróscopo proporcionará una medida de  $\omega$  que denotamos por  $\hat{\omega}$ . En realidad no será exactamente  $\omega$ , sino que estará contaminado por un cierto ruido blanco  $\nu$ :

$$\omega = \hat{\omega} + \nu$$

- Si intentamos estimar  $\theta$  (denotamos a la estimación  $\hat{\theta}$ ) directamente de  $\hat{\omega}$ , tendremos:

$$\dot{\hat{\theta}} = \hat{\omega}$$



## Ejemplo 1-D: propagación del error de un giróscopo

- El error de estimación  $\delta\theta = \theta - \hat{\theta}$  será el siguiente:

$$\delta\dot{\theta} = \omega - \hat{\omega} = \nu$$

- Suponiendo  $\nu$  ruido blanco (unidimensional) de varianza  $Q$  y  $\delta\theta(0) \approx N(0, P_0)$ , encontramos (usando la teoría de procesos expuesta en las primeras transparencias) que el error es un proceso estocástico gaussiano,  $\delta\theta \approx N(m, P)$ , donde:

$$\dot{m} = 0 \longrightarrow m = \delta\theta_0 = 0,$$

y

$$\dot{P} = \sigma_\nu^2 \longrightarrow P = P_0 + Qt$$

- Por tanto, aunque la media del error permanece fija en cero, la varianza del error crece linealmente con el tiempo y eventualmente se dispara, siendo por tanto este estimador inútil a medio plazo.



## Medida externa

- Supongamos que se tiene una medida externa adicional del ángulo. Suponemos que cada ciertos instantes discretos  $t = t_k$  se realiza una medida del ángulo  $\hat{\theta}(t_k)$ , que llamamos  $\hat{\theta}_k^{med}$  con algún otro dispositivo (que también tendrá un cierto error asociado, de forma que  $\theta_k = \hat{\theta}_k^{med} + \epsilon$ , donde  $\epsilon$  es ruido blanco, de varianza  $R$ ).
- Como el tiempo entre medidas puede ser grande, no es buena idea decir  $\hat{\theta}(t) = \hat{\theta}_k^{med}$  para  $t \in [t_k, t_{k+1})$ .
- Otra idea es resetear el estimador de las anteriores transparencias cuando se llega a  $t = t_k$ , es decir, combinar las medidas de la siguiente forma:
$$\hat{\theta} = \hat{\omega}, \quad \hat{\theta}(t_k) = \hat{\theta}_k^{med}, \quad t \in [t_k, t_{k+1}),$$
- Por tanto cada nueva medida externa se reinicia la condición inicial de la ecuación diferencial y se vuelve a integrar.
- Es fácil ver que el error obtenido de esta forma sería  $\delta\theta \approx N(m, P)$ , con  $m = 0$  y  $\dot{P} = Q$ , para  $t \in [t_k, t_{k+1})$ , con  $P(t_k) = R$ , luego  $P = R + Q(t - t_k)$ .



## Filtro de Kalman

- Con la anterior forma que el error sería máximo justo antes de una medida, obteniendo  $P = R + Q(t_{k+1} - t_k)$  en dicho instante.
- El problema es que se ha despreciado la estimación que daba la ecuación diferencial, cuando entre  $t_k$  y  $t_{k+1}$  no ha dado tiempo a que se el error crezca demasiado. La idea del filtro de Kalman es combinar la estimación de la ecuación diferencial justo antes de la medida externa, con la medida externa, de forma que la covarianza del error sea mínima.
- A la estimación justo antes de la medida se le llama estimación a priori  $\hat{\theta}_k^-$ .
- La estimación después de la medida (estimación a posteriori), se denota  $\hat{\theta}_k^+$  y se calcula como:

$$\hat{\theta}_k^+ = \hat{\theta}_k^- + K(\hat{\theta}_k^{med} - \hat{\theta}_k^-)$$

donde  $K$  es la ganancia de Kalman y el paréntesis es la diferencia entre la estimación que se tenía y la medida externa.



## Filtro de Kalman

- $K$  se calcula para minimizar la covarianza del error a posteriori.
- La covarianza a priori será  $P_k^-$ .
- A posteriori, calculando la covarianza de la ecuación de  $\theta_k^+$ :

$$P_k^+ = (1 - K)^2 P_k^- + K^2 R$$

- Derivando con respecto a  $K$  e igualando a cero:

$$0 = -2(1 - K)P_k^- + 2KR, \text{ luego } K = \frac{P_k^-}{P_k^- + R}.$$

- Por tanto la covarianza a posterior será, sustituyendo  $K$ :

$$P_k^+ = \frac{P_k^- R}{P_k^- + R}$$

- Se ve fácilmente que  $P_k^+$  es menor que  $R$  y menor que  $P_k^-$  (recordar que ambas son positivas): por tanto se ha conseguido mejorar tanto la estimación anterior como la que se tenía de la medida!



## Filtro de Kalman

- Resumiendo el algoritmo:
- Para  $t \in [t_k, t_{k+1})$ , se integra usando la medida de los giróscopos, partiendo de la última estimación a posteriori:

$$\dot{\hat{\theta}} = \hat{\omega}, \quad \theta(t_k) = \theta_k^+,$$

- También se propaga la covarianza del error:

$$\dot{P} = Q, \quad P(t_k) = P_k^+,$$

- Al llegar a  $t = t_{k+1}$ , se obtiene de estas ecuaciones  $\theta_{k+1}^- = \theta(t_{k+1})$  y  $P_{k+1}^-$ , y se obtiene una medida externa  $\hat{\theta}_{k+1}^{med}$ . Aplicamos el filtro de Kalman:

$$\hat{\theta}_{k+1}^+ = \hat{\theta}_{k+1}^- + K(\hat{\theta}_{k+1}^{med} - \hat{\theta}_{k+1}^-),$$

donde  $K = \frac{P_{k+1}^-}{P_{k+1}^- + R}$ , y también obtenemos  $P_{k+1}^+ = \frac{P_{k+1}^- R}{P_{k+1}^- + R}$ .

- Repetimos y volvemos a integrar las ecuaciones diferenciales hasta la nueva medida en  $t = t_{k+2}$ .



## Filtro de Kalman: casos extremos

- Si la medida del giróscopo es muy mala, es decir,  $Q$  es muy grande, entonces  $P_k^- \rightarrow \infty$ , y se puede ver que  $P_k^+ \rightarrow R$ ,  $K \rightarrow 1$ , y por tanto  $\theta_k^+ = \theta_k$  (es decir se coge la medida del sensor externo despreciando el resultado de integrar la ecuación diferencial).
- Si la medida del sensor externo es muy mala, es decir,  $R$  es muy grande, entonces se puede ver que  $P_k^+ \rightarrow P_k^-$ ,  $K \rightarrow 0$ , y por tanto  $\theta_k^+ = \theta_k^-$  (es decir se coge la estimación resultado de integrar la ecuación diferencial despreciando la medida del sensor externo ).
- Si resulta que  $P_k^- = R$ , es decir, la estimación a priori tiene el mismo error que es sensor externo, entonces se puede ver que  $P_k^+ = R/2$ ,  $K \rightarrow 1/2$ , y por tanto  $\theta_k^+ = \frac{\theta_k + \theta_k^-}{2}$  (es decir se toma la media entre la estimación resultado de integrar la ecuación diferencial y la medida del sensor externo ).



## Filtro de Kalman: otras consideraciones

- Se ha simplificado considerablemente el Filtro de Kalman al considerar un caso 1-D que es lineal. El caso N-D es similar pero más complejo ya que implica diversos productos e inversiones matriciales. No obstante conceptualmente es lo mismo: se integra la ecuación diferencial con los giróscopos y al obtener una medida externa se aplica el algoritmo de Kalman para ponderar entre la estimación a priori y la medida.
- Para aplicar el filtro de Kalman es suficiente una medida, pero cuantas más se tenga, mejor será el resultado.
- Si el sistema es no lineal hay que empezar por linealizarlo; el filtro resultante se denomina Filtro Extendido de Kalman. Para estimación de actitud esto es necesario ya que las ecuaciones diferenciales cinemáticas son siempre no-lineales.
- Para aeronaves y misiles se usa filtrado de Kalman para integrar las medidas de la IMU (giróscopos y acelerómetros) con medidas externas (GPS, antenas, magnetómetros) y así obtener no sólo la actitud sino también la posición.



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Vamos a desarrollar ahora el algoritmo del Filtro de Kalman en 3D para el siguiente caso:
  - 1 Se tienen giróscopos en los 3 ejes, de forma que se tiene una estimación de la velocidad angular  $\hat{\vec{\omega}}_{B/N}^B$ , que contendrá error, de forma que  $\vec{\omega}_{B/N}^B = \hat{\vec{\omega}}_{B/N}^B + \vec{v}$ , donde  $\vec{v}$  es ruido blanco de covarianza  $Q$ . Estas medidas se suponen continuas.
  - 2 Cada cierto tiempo se obtienen medidas de  $n$  direcciones en ejes cuerpo  $\hat{\vec{v}}_i^B$ , de forma que  $\vec{v}_i^B = C_N^B \vec{v}_i^N$  y  $\vec{v}_i^B = \hat{\vec{v}}_i^B + \vec{\epsilon}_i$  para  $i = 1, \dots, n$ .  $\vec{\epsilon}_i$  es ruido blanco de covarianza  $R_i$ .
- Se podrían realizar estimaciones con los giróscopos o con las medidas por separado. Con las medidas ya lo hemos visto (algoritmo Q) pero sólo se podría hacer si hubiera dos o más medidas tomadas a la vez.
- Con los giróscopos, la estimación sería  $\dot{\hat{q}} = \frac{1}{2} q \star q_{\hat{\vec{\omega}}}$ .



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Si uno usa  $\dot{\hat{q}} = \frac{1}{2} \hat{q} \star \hat{\omega}$ , ¿cuál es el error de estimación sabiendo que  $\hat{\omega}$  está corrompida con ruido blanco de covarianza  $Q$ ?
- No podemos aplicar la teoría general porque la ecuación es no-lineal!
- Recordamos los conceptos de cuaternión linealizado y ecuación cinemática linealizada:  $\dot{q} = \hat{q} \star \delta q$ , con

$$\delta q(\vec{a}) = \frac{1}{\sqrt{4 + \|\vec{a}\|^2}} \begin{bmatrix} 2 \\ \vec{a} \end{bmatrix}, \quad \dot{\vec{a}} \approx \vec{v} + \vec{a} \times \hat{\omega} = -\hat{\omega} \times \vec{a} + \vec{v}.$$

- Por tanto con el cuaternión de error podemos estimar cuanto error se está cometiendo, estudiando su covarianza  $P$ , la cual evolucionará de acuerdo con la teoría que hemos planteado antes:

$$\dot{P} = -\hat{\omega} \times P + P \hat{\omega} \times + Q, \quad P(0) = P_0$$



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Esta covarianza crece sin límite, lo cuál sólo podemos remediarlo tomando medidas. Si en un cierto instante tomamos  $n$  medidas, ¿cómo actualizamos el cuaternión estimado  $\hat{q}$ ?
- Una posibilidad sería descartar totalmente lo obtenido con los giróscopos, de forma que se resetea la estimación a un nuevo valor (sólo posible si se tienen dos o más medidas), usando el algoritmo Q. En tal caso la nueva covarianza sería la estudiada con el algoritmo Q.
- Una mejor idea sería aprovechar la estimación que ya teníamos y combinarlas de acuerdo al filtro de Kalman. Para ello hay que linealizar el proceso de medida. Del tema anterior,  
 $q_{\vec{v}_i^B} = q^* \star q_{\vec{v}_i^N} \star q$ , luego tenemos que  $q_{\vec{v}_i^N} = q \star q_{\vec{v}_i^B} \star q^*$  y  
 $\vec{v}_i^B = \hat{\vec{v}}_i^B + \vec{\epsilon}_i$ . Introduciendo el cuaternión de error:

$$q_{\vec{v}_i^N} = \hat{q} \star \delta q \star q_{\vec{v}_i^B} \star \delta q^* \star \hat{q}^* \approx \hat{q} \star \begin{bmatrix} 1 \\ \vec{a}/2 \end{bmatrix} \star \begin{bmatrix} 0 \\ \vec{v}_i^B \end{bmatrix} \star \begin{bmatrix} 1 \\ -\vec{a}/2 \end{bmatrix} \star \hat{q}^*$$



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Desarrollando a primer orden en  $\vec{a}$ :

$$\begin{aligned}
 q_{\vec{v}_i^N} &\approx \hat{q} \star \begin{bmatrix} -\vec{v}_i^B \cdot \vec{a}/2 \\ \vec{v}_i^B + \vec{a}/2 \times \vec{v}_i^B \end{bmatrix} \star \begin{bmatrix} 1 \\ -\vec{a}/2 \end{bmatrix} \star \hat{q}^* \\
 &= \hat{q} \star \begin{bmatrix} -\vec{v}_i^B \cdot \vec{a}/2 + \vec{a}/2 \cdot (\vec{v}_i^B + \vec{a}/2 \times \vec{v}_i^B) \\ (\vec{v}_i^B \cdot \vec{a}/2)\vec{a}/2 + \vec{v}_i^B + \vec{a}/2 \times \vec{v}_i^B - (\vec{v}_i^B + \vec{a}/2 \times \vec{v}_i^B) \times \vec{a}/2 \end{bmatrix} \star \hat{q}^* \\
 &= \hat{q} \star \begin{bmatrix} 0 \\ (\vec{v}_i^B \cdot \vec{a}/2)\vec{a}/2 + \vec{v}_i^B + \vec{a} \times \vec{v}_i^B + \vec{a}/2 \times (\vec{a}/2 \times \vec{v}_i^B) \end{bmatrix} \star \hat{q}^* \\
 &\approx \hat{q} \star \begin{bmatrix} 0 \\ \vec{v}_i^B + \vec{a} \times \vec{v}_i^B \end{bmatrix} \star \hat{q}^*
 \end{aligned}$$

- Llamamos  $\vec{z}_i$  a la discrepancia entre lo medido y lo esperado, es decir,  $\vec{z}_i = \hat{\vec{v}}_i^B - \hat{C}_N^B \vec{v}_i^N$ . Expresándolo con cuaterniones, será la parte vectorial de:

$$q_{\hat{\vec{v}}_i^B} - \hat{q}^* \star q_{\vec{v}_i^N} \star \hat{q} \approx q_{\hat{\vec{v}}_i^B} - \hat{q}^* \star \hat{q} \star \begin{bmatrix} 0 \\ \vec{v}_i^B + \vec{a} \times \vec{v}_i^B \end{bmatrix} \star \hat{q}^* \star \hat{q} \approx -q_{\vec{\epsilon}_i} + \begin{bmatrix} 0 \\ (\hat{\vec{v}}_i^B) \times \vec{a} \end{bmatrix}$$

- Por tanto tendremos  $n$  medidas del error de la forma  $\vec{z}_i = H_i \vec{a} - \vec{\epsilon}_i$ , donde  $H_i = (\hat{\vec{v}}_i^B)^\times$ . (NOTA: se tomarán sólo dos filas para evitar problemas de invertibilidad). La covarianza de esta medida será  $R_i$ .



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Llamemos al instante antes de la medida con el superíndice  $-$  y el posterior con  $+$ . De la integración teníamos por un lado  $\hat{q}^-$  con error  $a^-$  cuya media es  $E[\vec{a}^-] = 0$  y covarianza  $P^-$ .
- Con las medidas, supuestas independientes, formamos
$$\vec{z} = \begin{bmatrix} \vec{z}_1 \\ \vdots \\ \vec{z}_n \end{bmatrix}, H = \begin{bmatrix} H_1 \\ \vdots \\ H_n \end{bmatrix}, R = \begin{bmatrix} R_1 & & \\ & \ddots & \\ & & R_n \end{bmatrix}$$
- Usando las medidas, escribimos  $\vec{a}^+ = \vec{a}^- + K(\vec{z} - H\vec{a}^-)$ , pero puesto que en media  $\vec{a}^- = 0$ , simplemente tendremos:  $\vec{a}^+ = K\vec{z}$ , donde  $K$  es la ganancia de Kalman.
- La ganancia de Kalman (óptima) se calcula como  $K = P^- H^T (H P^- H^T + R)^{-1}$ . La covarianza se actualiza también como  $P^+ = P^- - K H P^-$ .
- Con  $\vec{a}^+$  corregimos  $\hat{q}:\hat{q}^+ = \hat{q}^- \star \delta q = \hat{q}^- \star \begin{bmatrix} 2 \\ \vec{a}^+ \end{bmatrix} \frac{1}{\sqrt{4 + \|\vec{a}^+\|^2}}$
- Este procedimiento se itera, integrando con los giróscopos y actualizando cada vez que hay medidas.



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Resumiendo el algoritmo. Datos iniciales:  $\hat{q}_0$ ,  $P_0$ ,  $Q$ ,  $R_i$ . En todo momento se tienen medidas de la velocidad angular  $\hat{\omega}$ . Ocasionalmente se tienen medidas  $\vec{z}_i$ .

- 1 Comenzar integrando las ecuaciones hasta que haya una medida, de forma que se estiman  $\hat{q}$  y  $P$ :

$$\dot{\hat{q}} = \frac{1}{2} q \star q_{\hat{\omega}}, \quad q(0) = q_0,$$

$$\dot{P} = -\hat{\omega}^\times P + P \hat{\omega}^\times + Q, \quad P(0) = P_0$$

- 2 Cuando en el instante  $t_k$  llegan una o más medidas, llamar  $\hat{q}^- = \hat{q}(t_k)$  y  $P^- = P(t_k)$ . Calcular  $\vec{z}$ ,  $H$ ,  $R$ . Calcular  $K = P^- H^T (H P^- H^T + R)^{-1}$ . Calcular  $\vec{a}^+ = K \vec{z}$ . Actualizar  $\hat{q}^+ = \hat{q}^- \star \delta q = \hat{q}^- \star \begin{bmatrix} 2 \\ \vec{a}^+ \end{bmatrix}$ ,  $P^+ = P^- - K H P^-$ .
- 3 Continuar integrando las ecuaciones a partir de las estimaciones actualizadas hasta que haya más medidas:

$$\dot{\hat{q}} = \frac{1}{2} q \star q_{\hat{\omega}}, \quad q(t_k) = q^+,$$

$$\dot{P} = -\hat{\omega}^\times P + P \hat{\omega}^\times + Q, \quad P(t_k) = P^+$$

- 4 Cuando haya nuevas medidas, iterar a partir de 2.



## Filtro Extendido Multiplicativo de Kalman (MEKF)

- Consideraciones adicionales:
  - Conviene renormalizar los cuaterniones estimados  $\hat{q}(t)$  si el módulo se alejara de la unidad.
  - Igualmente, la matriz de covarianzas  $P(t)$  debe ser simétrica en todo momento. Se puede "simetrizar" forzando  $P = 1/2(P + P^T)$ , o se puede simplemente calcular sólo la "mitad" de la matriz e imponer que la otra mitad es la traspuesta.
  - La ganancia de Kalman es óptima, pero está calculada para el problema linealizado. Si la estimación no está suficientemente cerca de la realidad, el filtro puede diverger.
  - Se pueden (y se deben) incluir los sesgos (bias) de los giróscopos en el estimador.
    - En la práctica es difícil conocer bien las matrices  $Q$  y  $R$ .
- Existen muchos otros algoritmos de filtrado, con sus ventajas e inconvenientes. El MEKF tiene la ventaja de ser sencillo en su uso y flexible a cualquier número de medidas, pero no es necesariamente el más preciso.

