Spacecraft Dynamics

Lesson 3: Attitude Determination. Errors.

Rafael Vázquez Valenzuela

Aerospace Engineering Department
Escuela Superior de Ingenieros, Universidad de Sevilla rvazquez1@us.es

June 30, 2022

Attitude determination

- Attitude determination is a process that estimates the present attitude by using sensors and applicable algorithms. It can be thought of as a "static" process that gives the picture of what the present attitude is.
- Attitude determination sensors, in general, determine a vector $\vec{v}$ in the body axes, this is, $\vec{v}^B$ (in fact they use "sensor axes" but the transformation to body axes should be known and it is implicitly applied). It is assumed that said vector is known in some reference axes (inertial axes or orbit axes), denoted as $\vec{v}^N$. As will be seen it is necessary to have two or more measurements of this kind to be able to solve the problem.
- In Lesson 6 we see sensors that from measurements of angular velocity $\vec{\omega}^B$ continuously determine the attitude (a more dynamic process that is typically referred to as attitude estimation).

Estimation from observations

- In general, consider we have n (2 or more) sensors that determine a vector $\vec{v}_i$, $i = 1, \ldots, n$, in body axes, this is, $\vec{v}_i^B$. The vector is assumed known in some reference axes (inertial axes or orbit axes, with respect to which we want to study the spacecraft attitude) and denoted in that frame as $\vec{v}_i^N$. Those are unit vectors since in principle only directions matter.
- Thus we have n equation written as $\vec{v}_i^B = C_N^B \vec{v}_i^N$ and we need to solve for C_N^B.
- To simplify write $\vec{W}_i = \vec{v}_i^B$, $\vec{V}_i = \vec{v}_i^N$, $A = C_N^B$. Thus, we have n equations $\vec{W}_i = A\vec{V}_i$ and need to solve for A.
- These vectors will contain some errors.
- If $n = 2$ there a simple method that can be applied known as TRIAD. We'll see other more general methods for $n \geq 2$.
- Question: what conditions would the measurements/references verify if they are exact??

TRIAD Method

- Start from two observations related to the references through the DCM: $\vec{W}_1 = A\vec{V}_1$ and $\vec{W}_2 = A\vec{V}_2$
- Define the following vectors: $\vec{r}_1 = \vec{V}_1$, $\vec{r}_2 = \frac{\vec{V}_1 \times \vec{V}_2}{|\vec{V}_1 \times \vec{V}_2|}$, and $\vec{r}_3 = \frac{\vec{V}_1 \times \vec{r}_2}{|\vec{V}_1 \times \vec{r}_2|}$. Similarly: $\vec{s}_1 = \vec{W}_1$, $\vec{s}_2 = \frac{\vec{W}_1 \times \vec{W}_2}{|\vec{W}_1 \times \vec{W}_2|}$, and $\vec{s}_3 = \frac{\vec{W}_1 \times \vec{s}_2}{|\vec{W}_1 \times \vec{s}_2|}$. It is rather obvious that one should have now: $\vec{s}_1 = A\vec{r}_1$, $\vec{s}_2 = A\vec{r}_2$, and $\vec{s}_3 = A\vec{r}_3$.
- Construct now the matrices $M_{ref} = [\vec{r}_1 \; \vec{r}_2 \; \vec{r}_3]$ and $M_{obs} = [\vec{s}_1 \; \vec{s}_2 \; \vec{s}_3]$. It holds that $M_{obs} = AM_{ref}$. In addition, the columns of M_{ref} are orthonormal between them. Thus, M_{ref} is invertible (and orthogonal!). Therefore we can solve for A as $A = M_{obs}M_{ref}^T$.
- Notice that the method is not symmetric, as the measurement labelled as 1 is given more importance. In practice, A will not be the exact DCM matrix due to errors in the sensors. Thus, one should use the "best" measurement as first.

Wahba's Problem

- Consider now n measures satisfying $\vec{W}_i = A\vec{V}_i$. We pose the problem as a least squares minimization problem.
- Define the function $L(A) = \frac{1}{2}\sum_{i=1}^{n} a_i|\vec{W}_i - A\vec{V}_i|^2$, where a_i are the weights given to each measurement (verifying $\sum_{i=1}^{n} a_i = 1$) and pose the mathematical objective of finding A (orthogonal) such $L(A)$ is minimized. In the literature this is known as **"Wahba's Problem"**.
- Since operating

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

one has

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

where $g(A) = \sum_{i=1}^{n} a_i \vec{W}_i^T A\vec{V}_i$. Minimizing $L(A)$ is thus equivalente to maximizing $g(A)$ (and notice $g(A) \leq 1!$).

Davenport's q method

- Writing A as a function of q by using Euler-Rodrigues $(A = (q_0^2 - \vec{q}^T\vec{q})\mathrm{I} + 2\vec{q}\vec{q}^T - 2q_0\vec{q}^\times)$ we reach

$$g(A) = \sum_{i=1}^{n} a_i \vec{W}_i^T (q_0^2 - \vec{q}^T\vec{q})\vec{V}_i + 2\sum_{i=1}^{n} a_i \vec{W}_i^T \vec{q}\vec{q}^T\vec{V}_i - 2\sum_{i=1}^{n} a_i \vec{W}_i^T q_0\vec{q}^\times \vec{V}_i$$

- Develop now each term trying to reach a bilineal form $g(q) = q^T K q$:
  - Starting with the second term

$$2\sum_{i=1}^{n} a_i \vec{W}_i^T \vec{q}\vec{q}^T\vec{V}_i = 2\sum_{i=1}^{n} a_i \vec{q}^T \vec{W}_i \vec{V}_i^T \vec{q} = 2\vec{q}^T B\vec{q} = \vec{q}^T (B + B^T)\,\vec{q}$$

where $B = \sum_{i=1}^{n} a_i \vec{W}_i \vec{V}_i^T$.
  - The first term can be written as

$$\sum_{i=1}^{n} a_i \vec{W}_i^T (q_0^2 - \vec{q}^T\vec{q})\vec{V}_i = (q_0^2 - \vec{q}^T\vec{q})\sum_{i=1}^{n} a_i \vec{W}_i^T\vec{V}_i = q_0\sigma q_0 - \vec{q}^T(\sigma\mathrm{I})\vec{q}$$

where $\sigma = \sum_{i=1}^{n} a_i \vec{W}_i^T \vec{V}_i = \mathrm{Tr}(B)$.

Davenport's q method

- Finally, the last term can be expressed as:

$$-2\sum_{i=1}^{n} a_i \vec{W}_i^T q_0 \vec{q}^\times \vec{V}_i = 2\sum_{i=1}^{n} a_i \vec{W}_i^T q_0 \vec{V}_i^\times \vec{q} = 2q_0 \vec{z}^T \vec{q} = q_0 \vec{z}^T \vec{q} + \vec{q}^T \vec{z} q_0$$

  where $\vec{z}^T = \sum_{i=1}^{n} a_i \vec{W}_i^T \vec{V}_i^\times$, hence $\vec{z} = -\sum_{i=1}^{n} a_i \vec{V}_i^\times \vec{W}_i$.
- One has $(\vec{a}^\times \vec{b})^\times = \vec{b}\vec{a}^T - \vec{a}\vec{b}^T$, what can be shown from the identity $(\vec{a} \times \vec{b}) \times \vec{c}$. Observe then that

$$\vec{z}^\times = -\sum_{i=1}^{n} a_i (\vec{V}_i^\times \vec{W}_i)^\times = \sum_{i=1}^{n} a_i \vec{V}_i \vec{W}_i^T - \sum_{i=1}^{n} a_i \vec{W}_i \vec{V}_i^T = B^T - B$$

Davenport's q method

- Thus, the function g is expressed in terms of the quaternion as

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

where the matrix K can be found from the coefficients of a newly defined matrix in terms of weights, measurements and references $B = \sum_{i=1}^{n} a_i \vec{W}_i \vec{V}_i^T$, as follows

$$\begin{aligned} \sigma &= \mathrm{Tr}(B), \\ S &= B + B^T, \\ \vec{z}^{\times} &= B^T - B \end{aligned}$$

being K a 4×4 matrix equal to

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

Davenport's q method

- Thus, the problem is now reduced to finding q (attitude quaternion, this is, a norm 1 vector of four components) such that $g(q) = q^T K q$ is maximized.
- To solve a multivariable maximization problem with constraints ($q^T q = 1$) one can use Lagrange's multipliers:

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

- Taking derivative w.r.t. q and setting it to zero: $\frac{\partial H}{\partial q} = 2q^T K - 2\lambda q^T = 0 \quad \longrightarrow \quad Kq = \lambda q.$
- Thus λ must be an eigenvalue of K and q the associated eigenvector of modulus 1 (there are two, but of opposing signs, thus representing the same attitude). To find which eigenvalue, replace the solution in $g(q)$: $g(q) = q^T K q = q^T \lambda q = \lambda$
- Therefore, the maximum attained at the critical point is equal to the eigenvalue and the solution will be the eigenvector (of modulus 1) associated to the maximum eigenvalue.

The QUEST method

- Davenport's q method reduces the attitude determination problem to an eigenvalue/eigenvector problem, however this algebraic method might be problematic to solve on a satellite, depending on computational resources available onboard.
- In 1978 the QUEST (QUaternion ESTimator) method was developed to avoid the computational burden.
- The idea is to rewrite $Kq = \lambda q$ in terms of the K matrix:

$$\begin{bmatrix} \sigma & \vec{z}^T \\ \vec{z} & S - \sigma \mathrm{Id} \end{bmatrix} \begin{bmatrix} q_0 \\ \vec{q} \end{bmatrix} = \lambda \begin{bmatrix} q_0 \\ \vec{q} \end{bmatrix}$$

- Therefore two equations can be extracted.

$$\sigma q_0 + \vec{z}^T \vec{q} = \lambda q_0, \quad q_0 \vec{z} + S\vec{q} - \sigma \vec{q} = \lambda \vec{q}$$

- Remembering Gibb's vector $\vec{g} = \frac{\vec{q}}{q_0}$, one can manipulate the second equation reaching

$$\vec{z} + [S - (\sigma + \lambda)\mathrm{I}]\, \vec{g} = 0$$

The QUEST Method

- Then $\vec{g} = [(\sigma + \lambda)\mathrm{I} - S]^{-1}\,\vec{z}$ (but we don't know λ, the maximum eigenvalue)
- A first approximation is to take $\lambda \approx 1$ (which would be the value if the measurements were without error). Then $\vec{g} = [(1 + \sigma)\mathrm{I} - S]^{-1}\,\vec{z}$
- A better approximation is to find an explicit expression for the maximum eigenvalue by finding the roots of the characteristic equation of K, which is:

$$\lambda^4 - (a + b)\lambda^2 - c\lambda + (ab + c\sigma - d) = 0$$

- Where the coefficients are

$$\begin{aligned} a &= \sigma - \mathrm{Tr}[\mathrm{adj}(S)], \\ b &= \sigma - \vec{z}^T\vec{z}, \\ c &= \det[S] + \vec{z}^T S\vec{z}, \\ d &= \vec{z}^T S^2\vec{z}. \end{aligned}$$

Errors in attitude determination

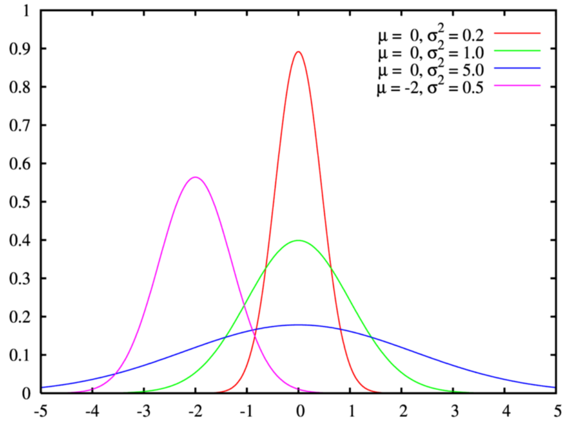
- Errors are, by definition, unknown. Since, if they were known, they would not be errors anymore!
- However, it is important to characterize errors in some way.
- The science that deals with unknowns is statistics (and its associated math field, probability).
- Engineers have to know about statistics, since it can be applied to many fields. Here, we give a refresher for some concepts necessary for estimating errors in attitude determination.
- We will always use normal distributions.
- We go from sensor errors (typically given by their technical specifications) to errors in attitude determination: propagation of uncertainty.

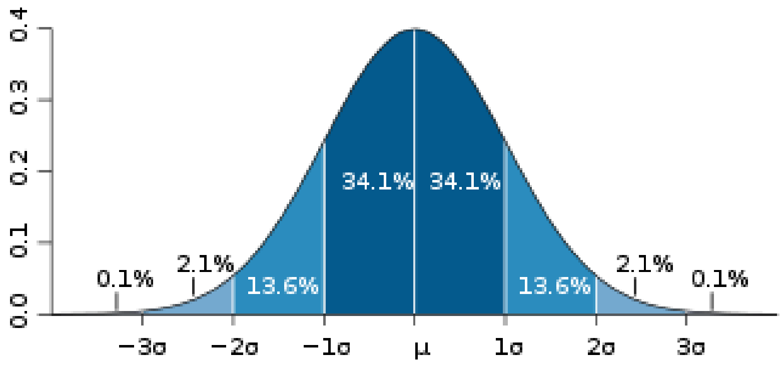
1-D Continuous RandomVariables

- Let $X \in \mathbb{R}$ be a random continuous variable.
- Remember that the cumulative distribution function (CDF) $F(x)$ is the probability that $X \leq x$, which is written as $F(x) = P(X \leq x)$.
- The CDF is computed from the probability density function (PDF) $f(x)$: $F(x) = \int_{-\infty}^{x} f(y)dy$.
- One defines the operator "mathematical expectation" acting over the function $g(x)$ as $E[g(X)] = \int_{-\infty}^{\infty} g(y)f(y)dy$. It is a linear operator: $E[\alpha_1 g_1(X) + \alpha_2 g_2(X)] = \alpha_1 E[g_1(X)] + \alpha_2 E[g_2(X)]$. Two importan examples are:
  - Mean: $m(X) = E[X] = \int_{-\infty}^{\infty} yf(y)dy$.
  - Variance: $V(X) = E[(X - m(X))^2] = E[X^2] - (E[X])^2$ (non-negative).
  - The typical deviation σ is the square root of the variance $\sigma = \sqrt{V(X)}$ to make it have the same units as the mean.
- Does it make sense for errors to have nonzero mean?

Normal (Gaussian) distribution I

- It is the most commonly used distribution in statistics. One writes $X \sim N(m, \sigma^2)$ and its PDF is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{Exp}\left(-\frac{(x-m)^2}{2\sigma^2}\right)$.
- Confidence intervals: if $X \sim N(m, \sigma^2)$ then:
  - 1-σ interval: $P(X \in [m-\sigma, m+\sigma]) = 68.3\%$.
  - 2-σ interval: $P(X \in [m-2\sigma, m+2\sigma]) = 95.45\%$.
  - 3-σ interval: $P(X \in [m-3\sigma, m+3\sigma]) = 99.74\%$.





Normal (Gaussian) distribution II

- The central limit theorem shows that the sum of independent random variables (with any kind of distribution), tends (in average) to a normal distribution. Since large-scale errors come from the sum and accumulation of many small-scale errors (think for example about temperature fluctuations), this justifies using normal distributions as a good model for errors.
- An important property of a normal distribution is that the sum of independent normals is again normal, this is, if $X \sim N(m_x, \sigma_x^2)$ and $Y \sim N(m_y, \sigma_y^2)$, and they are independent, then $Z = X + Y$ is distributed as $Z \sim N(m_x + m_y, \sigma_x^2 + \sigma_y^2)$.
- Therefore $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$, this is, the typical deviation of the sum of errors is the square root of the sum of squares of the typical deviation of errors.
- This rule is known as Root-Sum-of-Squares (RSS) and it is of high importance when dealing with accumulated errors.

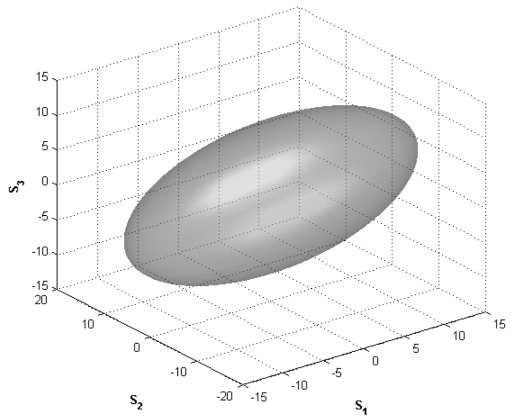
Multivariate Continuous Random Variables

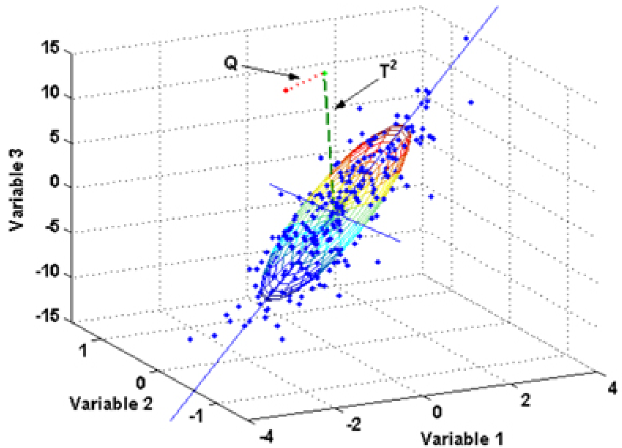
- Let $\vec{X} \in \mathbb{R}^n$ be a multivariate continuous random variables.
- Each component of $\vec{X}$ follows a 1-D distribution (i.e. is a 1-D random variable).
- Following the 1-D case, we now define a *joint* CDF that is computed from a joint PDF $f(\vec{x})$.
- Similarly $E[g(\vec{X})] = \int_{\mathbb{R}^n} g(\vec{y})f(\vec{y})dy$. Important cases:
  - Mean: $\vec{m}(\vec{X}) = E[\vec{X}] = \int_{\mathbb{R}^n} \vec{y}f(\vec{y})dy$.
  - Covariance: $Cov(\vec{X}) = E[(\vec{X} - m(\vec{X}))(\vec{X} - m(\vec{X}))^T] = \Sigma$. A symmetric, non-negative definite matrix. The values of its diagonal represent the variance the corresponding component of $\vec{X}$, whereas off-diagonal coefficients represent the correlation between two components of $\vec{X}$. One has $\Sigma = E[(\vec{X}\vec{X}^T] - m(\vec{X})m(\vec{X})^T$.
- For instance for $n = 3$ and writing $\vec{X} = [X, Y, Z]$:

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

Multivariate normal distribution I

- One writes $\vec{X} \sim N_n(\vec{m}, \Sigma)$ and its PDF is
$f(\vec{x}) = \frac{1}{\mathrm{Det}(\Sigma)(2\pi)^{n/2}} \mathrm{Exp}\left(-\frac{1}{2}(\vec{x}-\vec{m})^T \Sigma^{-1}(\vec{x}-\vec{m})\right)$.
- Confidence intervals become regions in $\mathbb{R}^n$, defined by
$P(\vec{X} \in \Omega) = P_\Omega$.
- The shape of these regions is a multidimensional ellipsoid described by $(\vec{x}-\vec{m})^T \Sigma^{-1}(\vec{x}-\vec{m}) = d^2$, where d depends on P_Ω. The size of the eigenvalues of Σ determines the size of the ellipsoid, whereas the direction of the ellipsoid axes is given by the eigenvectors of Σ.





Multivariate normal distribution II

- A classical example from aerial navigation or orbital mechanics, one can describe an aircraft/spacecraft position in some axes as $\delta\vec{r} = [\delta x \ \delta y \ \delta z]^T$, as a multivariate normal with $n = 3$, with mean zero (centered in the expected position of the vehicle) and covariance matrix

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

- Then one can visualize the movement of the vehicle with the movement of the whole ellipsoid, representing a region (tube) where the vehicle can be found with some degree of certainty.
- Property: If $\vec{X} \sim N_n(\vec{m}_x, \Sigma_x)$ and $\vec{Y} \sim N_n(\vec{m}_y, \Sigma_y)$ and they are independent, then if $\vec{Z} = \vec{X} + \vec{Y}$ it follows that $\vec{Z} \sim N_n(\vec{m}_x + \vec{m}_y, \Sigma_x + \Sigma_y)$.
- Similarly if $A\vec{X} + \vec{b}$ where A and b are non-random (known) it follows that $A\vec{X} + \vec{b} \sim N_n(A\vec{m}_x + \vec{b}, A\Sigma_x A^T)$.

Errors in attitude determination

- How can one characterize attitude errors?
- It will depend on the chosen attitude representation.
- For instance if one chooses quaternions, then one could use the quaternion error, parameterized $\delta q(\vec{a})$ and give a multivariate distribution for $\vec{a}$. Typically with zero mean and some covariance. Then the approximate attitude $\hat{q}$ is related to the real attitude q as in Lesson 2: $\hat{q} = q \star \delta q$ where

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

- If one uses the DCM, it is required to find a way to represent some kind of "DCM error".
- It does not make sense to use a 9-dimensional distribution function to characterize the error of each component since attitude does have 3 degrees of freedom, as we know.
- Since errors are (or should be) small, we next characterize DMC errors with an approximation for "small" DMC.

DCM for small angles I

- Let A and B be two reference frames related as follows

$$A \xrightarrow[x^A]{d\theta_1} S_1 \xrightarrow[y^{S_1}]{d\theta_2} S_2 \xrightarrow[z^{S_2}]{d\theta_3} B$$

  where we assume that $d\theta_i$ are small angles, so we can make the approximations $\cos d\theta_i \simeq 1$ and $\sin d\theta_i \simeq d\theta_i$.

- Writting the DCMs taking into account the approximations:

$$C_A^{S_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & d\theta_1 \\ 0 & -d\theta_1 & 1 \end{bmatrix}, \; C_{S_1}^{S_2} = \begin{bmatrix} 1 & 0 & -d\theta_2 \\ 0 & 1 & 0 \\ d\theta_2 & 0 & 1 \end{bmatrix}, \; C_{S_2}^{B} = \begin{bmatrix} 1 & d\theta_3 & 0 \\ -d\theta_3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- Then, since $C_A^B = C_{S_2}^B C_{S_1}^{S_2} C_A^{S_1}$, and neglecting all double products of angles (i.e. $d\theta_i d\theta_j \simeq 0$), one gets:

$$C_A^B = \begin{bmatrix} 1 & d\theta_3 & -d\theta_2 \\ -d\theta_3 & 1 & d\theta_1 \\ d\theta_2 & -d\theta_1 & 1 \end{bmatrix} = \mathrm{Id} - \begin{bmatrix} 0 & -d\theta_3 & d\theta_2 \\ d\theta_3 & 0 & -d\theta_1 \\ -d\theta_2 & d\theta_1 & 0 \end{bmatrix} = \mathrm{Id} - d\vec{\theta}^{\times},$$

DCM for small angles II

- In the previous slides the definition $d\vec{\theta} = [d\theta_1 \; d\theta_2 \; d\theta_3]^T$ was made, and the matrix

$$d\vec{\theta}^{\times} = \begin{bmatrix} 0 & -d\theta_3 & d\theta_2 \\ d\theta_3 & 0 & -d\theta_1 \\ -d\theta_2 & d\theta_1 & 0 \end{bmatrix},$$

  is the result of the operator $\times$ as was defined in Lesson 2.
- Notice that under these hypothesis (small angles) it does not matter the order of rotations and the angles add up, however not all sets of Euler angles could be used since no axes can be be repeated (meaning: 1-2-3 o 3-2-1 or any similar set works, but 1-2-1 would not).
- Exercise: work out the (very simple!) relationship between the small angles vector and the vector $\vec{a}$ used in quaternion errors by using Euler's axis and angle.

Error of a DCM

- To model errors for a DCM we use the "small angles vector" just defined, which will be randomly distributed.
- Denote $\hat{C}_N^B$ the matrix with errors (or actually $\hat{C}_N^B = C_n^{\hat{B}}$), where:

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

- Then $C_N^{\hat{B}} = C_B^{\hat{B}} C_N^B$ and thus $C_N^B = C_{\hat{B}}^B C_N^{\hat{B}}$, and we define $\delta C_N^B = C_N^B - \hat{C}_N^B = C_{\hat{B}}^B \hat{C}_N^B - \hat{C}_N^B = (C_{\hat{B}}^B - \mathrm{Id})\hat{C}_N^B$.
- Assuming $\delta\vec{\phi} = [\delta\phi_x\ \delta\phi_y\ \delta\phi_z]^T$ are small, one has $C_B^{\hat{B}} = \mathrm{Id} - \delta\vec{\phi}^\times$ (and $C_{\hat{B}}^B = \mathrm{Id} + \delta\vec{\phi}^\times$).
- Then the relationship between the "error matrix" δC_N^B and $\delta\vec{\phi}$ is $\delta C_N^B = (\mathrm{Id} + \delta\vec{\phi}^\times - \mathrm{Id})\hat{C}_N^B = \delta\vec{\phi}^\times \hat{C}_N^B$.And one has $C_N^B = (\mathrm{Id} + \delta\vec{\phi}^\times)\hat{C}_N^B$.

Covariance matrix for TRIAD

- For TRIAD, one can model the error as a small angles vector $\delta\vec{\phi}$ given by a multivariate normal with zero mean and covariance $P_{\phi\phi}$. One can prove:

$$P_{\phi\phi} = \sigma_1^2 \mathrm{Id} + \frac{1}{|\vec{W}_1 \times \vec{W}_2|^2}\left((\sigma_2^2 - \sigma_1^2) W_1 W_1^T + \sigma_1^2 (W_1^T W_2)(W_1 W_2^T + W_2 W_1^T)\right)$$

  where σ_1 represents the angular error (given as typical deviation) of the first measurement and σ_2 the error of the second measurement.
- Notice, as expected, that the first measurement has more influence on the final error.
- If the measurements are orthogonal, then:

$$P_{\phi\phi} = \sigma_1^2 \mathrm{Id} + (\sigma_2^2 - \sigma_1^2) W_1 W_1^T$$

- Imagine for instance if W_1 is the x axis, then this results in $P_{\phi\phi}$ diagonal, with the (1,1) entry as σ_2^2 and the other diagonal coefficients as σ_1^2: Can you interpret this?

Covariance matrix for q

- Now $\vec{a}$ represents the attitude error (via $\delta q(\vec{a})$) and therefore we model $\vec{a}$ as a multivariate distributed vector with zero mean and covariance matrix P_a.
- In the q algorithm each measurement has an error represented by its variance σ_i^2. The global error of q depends on the chosen weights and one can prove the following relationship

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

- A good rule of thumb for a_i is make it proportional to the inverse of the variances σ_i^2, however since the a_i's add up to 1, one chooses $a_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}}$ so $P_a = \left[\sum_{j=1}^{n} \frac{1}{\sigma_j^2} \mathrm{Id} - \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \vec{W}_i \vec{W}_i^T \right]^{-1}$.
- Note that $\left[\mathrm{Id} - \sum_{i=1}^{n} a_i \vec{W}_i \vec{W}_i^T \right]$ should be invertible.
- Exercise: consider the particular case analyzed for TRIAD with equal weights and compare.