Ellipsoid Fit

To calibrate the magnetometer, one method is to fit an ellipsoid to the measured data and adjust the ellipsoid into a sphere. So first going to talk about how to fit an ellipsoid with a set of measurements.

A quadratic surface (such as an ellipsoid, cone, cylinders) can be constructed by the following equation:

$$ax^2 + by^2 + cz^2 + 2fyz + 2gxz + 2hxy + 2px + 2qy + 2rz + d = 0$$

and in matrix form, we could rewrite the above equation into

$$\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} + 2\boldsymbol{x}^T \boldsymbol{n} + d = 0$$

where

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \boldsymbol{M} = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \qquad \boldsymbol{n} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}$$

So to construct an ellipsoid we need to know the values of $\boldsymbol{M}, \boldsymbol{n}$ and d. If we have a set of measured data from the magnetometer, then we could use the method from this paper by Qingde Li and John G. Griffiths which best fits a set of data into an ellipsoid using a least squared method. Then we could get the estimated value $\hat{\boldsymbol{M}}, \hat{\boldsymbol{n}}$ and $\hat{d}$. A quick summary of this process is mentioned at the end.

Magnetometer Calibration

Now we know how to get the best fit ellipsoid, we can calibrate the magnetometer to correct the offset from hard-iron biases and distortion caused by the soft-iron biases.

Define $\boldsymbol{h}$ as the true magnetic field, and with distortion we can get the measured magnetic value $\boldsymbol{h}_m$ with the following:

$$\boldsymbol{h}_m = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{b}$$

where $\boldsymbol{b}$ is the offset caused by the hard-iron bias, and $\boldsymbol{A}$ is the effect from soft-iron bias. But the magnetometer is giving you the measured magnetic value $\boldsymbol{h}_m$ so to get the true magnetic value $\boldsymbol{h}$ we have

$$\boldsymbol{h} = \boldsymbol{A}^{-1}(\boldsymbol{h}_m - \boldsymbol{b})$$

Define the magnitude of the true magnetic field as $\mathcal{F}$ such that we would get $\boldsymbol{h}^T\boldsymbol{h} = \mathcal{F}^2$. Then if we insert the above equation into $\boldsymbol{h}^T\boldsymbol{h} = \mathcal{F}^2$ we get

$$\boldsymbol{h}_m^T \boldsymbol{A}^{-T} \boldsymbol{A}^{-1} \boldsymbol{h}_m - 2\boldsymbol{h}_m^T \boldsymbol{A}^{-T} \boldsymbol{A}^{-1} \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{A}^{-T} \boldsymbol{A}^{-1} b - \mathcal{F}^2 = 0$$

we can put the above into matrix form which is

$$\boldsymbol{h}_m^T \boldsymbol{M} \boldsymbol{h}_m + 2\boldsymbol{h}_m^T \boldsymbol{n} + d = 0$$

where

$$\boldsymbol{M} = \boldsymbol{A}^{-T}\boldsymbol{A}^{-1} \qquad \boldsymbol{n} = -\boldsymbol{A}^{-T}\boldsymbol{A}^{-1}\boldsymbol{b} \qquad d = \boldsymbol{b}^T\boldsymbol{A}^{-T}\boldsymbol{A}^{-1}\boldsymbol{b} - \mathcal{F}^2$$

Note that the equation defined above is similar to the matrix form of the quadratic surface equation. And after we get the estimated $\hat{\boldsymbol{M}}, \hat{\boldsymbol{n}}$, and $\hat{d}$ from the ellipsoid fit method by Qingde Li and John G. Griffiths, we can get an estimated value for A^{-1} and $\boldsymbol{b}$. After some math manipulation, the estimated values for $\boldsymbol{A}$ and $\boldsymbol{b}$ are:

$$\hat{\boldsymbol{b}} = -\hat{\boldsymbol{M}}^{-1}\hat{\boldsymbol{n}}$$

$$\hat{\boldsymbol{A}}^{-1} = \frac{\mathcal{F}}{\sqrt{\hat{\boldsymbol{n}}^T \hat{\boldsymbol{M}}^{-1} \hat{\boldsymbol{n}} - \hat{d}}} \hat{\boldsymbol{M}}^{\frac{1}{2}}$$

and with these estimated values, we can get the magnetic field $\boldsymbol{h}$

$$\boldsymbol{h} = \hat{\boldsymbol{A}}^{-1}(\boldsymbol{h}_m - \hat{\boldsymbol{b}})$$

Note that we need to know the value $\mathcal{F}$ in order to get the proper scaling of $\hat{\boldsymbol{A}}^{-1}$. This can be determined by using a magnetic field model such as those provided by NOAA. (I have not implemented this part).

Test Result

Fig. 1 shows an example of how the data set looks like before and after calibrating. In the image, we have a planar view of the measurements. The blue dots are the measurements only looking at the X and Y values, the green is only looking at the Z and X values, and the orange is only looking at the Y and Z values. On the left, we have the uncalibrated measurement and we could that the shape are ellipsoidal and isn't centered around the origin. An ideal measurement would have a circular shape and centered around the origin, as seen in the right plot which is the result after calibrating.

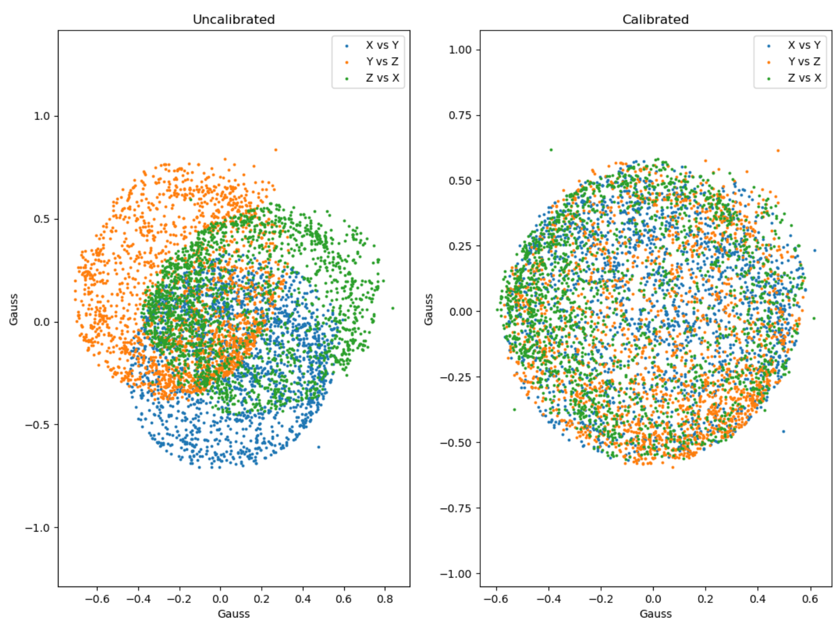

Figure 1: Before and after calibration

Another calibration result from a 3D perspective using a different dataset is shown in Fig. 2. The blue dots are the uncalibrated magnetic field, and you could see that the data set has an ellipsoidal surface. It's hard to tell from this image, but the ellipsoid isn't centered around the origin. Now after calibrating the measurement, we get the orange data set which has a spherical surface and centered around the origin.

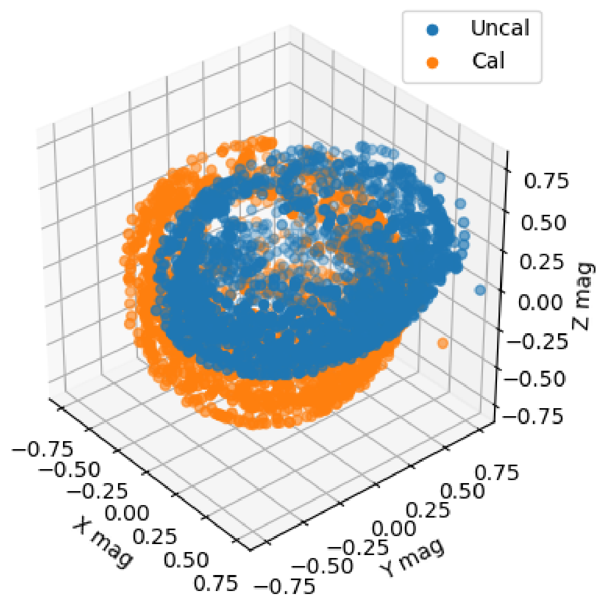

Figure 2: Another before and after calibration

Brief Summary on Ellipsoid Fitting

To fit an ellipsoid to a set of data, we first need to gather the data. So gather enough data so that the magnetometer see all measurement in each axis (hard to explain). Say we gathered N measurements, then for each measurement $\{x_1, y_i, z_i\}$ where $i \in 0, 1, 2, ..., N-1$ we will define a column vector X_i as

$$X_i = \{x_i^2, y_i^2, z_i^2, 2y_iz_i, 2x_iz_i, 2x_iy_i, 2x_i, 2y_i, 2z_i, 1\}^T$$

Then we create a matrix D which is defined as follow:

$$D = \{X_0, X_1, X_2, ...X_N\}$$

Also define the following matrix:

$$C = \begin{bmatrix} -1 & \frac{k}{2}-1 & \frac{k}{2}-1 & 0 & 0 & 0 \\ \frac{k}{2}-1 & -1 & \frac{k}{2}-1 & 0 & 0 & 0 \\ \frac{k}{2}-1 & \frac{k}{2}-1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k & 0 & 0 \\ 0 & 0 & 0 & 0 & -k & 0 \\ 0 & 0 & 0 & 0 & 0 & -k \end{bmatrix}$$

For our implementation, we set $k = 5$ which got us a good result. What we understand about what k does is it's a tuning parameter when the data set is noisy.

After computing the matrix D, we need to get matrix S which is a 10x10 matrix and defined as follows:

$$\begin{aligned} S &= D^T D \\ &= \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \end{aligned}$$

where S_{11} is a 6x6 matrix, S_{12} is a 6x4 matrix, and S_{22} is a 4x4 matrix. Note that S is a symmetric matrix so $S_{12}^T = S_{21}$.

The next step is to get find v_1 and λ from the following, which is an eigenvalue problem:

$$C^{-1}(S_{11} - S_{12}S_{22}^{-1}S_{21})v_1 = \lambda v_1$$

If the first element in v_1 is negative, then set $v_1 = -v_1$. After that we need to get v_2:

$$v_2 = -S_{22}^{-1}S_{21}v_1$$

Then finally we can define the estimated $\hat{\boldsymbol{M}}$, $\hat{\boldsymbol{n}}$, and $\hat{d}$:

$$\hat{\boldsymbol{M}} = \begin{bmatrix} v_1(0) & v_1(5) & v_1(4) \\ v_1(5) & v_1(1) & v_1(3) \\ v_1(4) & v_1(3) & v_1(2) \end{bmatrix}$$

$$\hat{\boldsymbol{n}} = \begin{bmatrix} v_2(0) \\ v_2(1) \\ v_2(2) \end{bmatrix}$$

$$\hat{d} = v_2(3)$$

and with this we got the parameters to determine the surface of the ellipsoid.

Calibrating with CubeSAT

As you noticed, there's a lot of linear algebra in this calibration. Currently, I did not find any linear algebra operation for the FreeRTOS. In Linux, there's a GNU GSL library that can do those operation (except for square of a matrix, which is available in example 6 in the magnetometer repository, under RPi). For the actual calibration, when we calibrate the magnetometer with the CubeSAT, one thing we could do is to gather the data, store it into a SD card, then transfer those data into a device with Linux OS, and then calibrate the data similar to example 6. After getting the matrix A and b, we can manually input those values into the actual CubeSAT. Some concern is that we can only calibrate the magnetometer initially. We can't calibrate the magnetometer during it's mission, but do we need to worry about this?