Module 1 | Lesson 1 (Part 2)

Weighted Least Squares

Weighted Least Squares

By the end of this video, you will be able to…

- Derive the weighted least squares criterion given varying measurement noise variance

- Compare weighted least squares to ordinary least squares

Method of *Weighted* Least Squares

- Suppose we take measurements with multiple multimeters, some of which are better than others

vs.

Method of *Weighted* Least Squares

Consider the general linear measurement model for m measurements and n unknowns:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$$

In regular least squares, we implicitly assumed that each noise term was of equal variance:

$$\mathbb{E}[v_i^2] = \sigma^2 \quad (i = 1, \ldots, m) \qquad\qquad \mathbf{R} = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix}$$

Method of *Weighted* Least Squares

If we assume each noise term is independent, but of different variance,

$$\mathbb{E}[v_i^2] = \sigma_i^2, \quad (i = 1, \ldots, m) \qquad \mathbf{R} = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_m^2 \end{bmatrix}$$

Then we can define a *weighted* least squares criterion as:

$$\begin{aligned} \mathscr{L}_{\mathrm{WLS}}(\mathbf{x}) &= \mathbf{e}^T \mathbf{R}^{-1} \mathbf{e} \\ &= \frac{e_1^2}{\sigma_1^2} + \frac{e_2^2}{\sigma_2^2} + \ldots + \frac{e_m^2}{\sigma_m^2} \end{aligned}$$

where

$$\begin{bmatrix} e_1 \\ \vdots \\ e_m \end{bmatrix} = \mathbf{e} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} - \mathbf{H} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

The higher the expected noise, the lower the weight we place on the measurement.

Re-deriving regular least squares

What happens if all of our variances are the same?

$$\mathscr{L}_{\text{WLS}}(\mathbf{x}) = \frac{e_1^2}{\sigma^2} + \frac{e_2^2}{\sigma^2} + \ldots + \frac{e_m^2}{\sigma^2}$$
$$= \frac{1}{\sigma^2}(e_1^2 + \ldots + e_m^2)$$

Since our variance is constant, it will not affect our final estimate! *This results in the same estimate as regular least squares:*

$$\sigma_1^2 = \sigma_2^2 = \ldots \sigma_m^2 = \sigma^2 \rightarrow \quad \hat{\mathbf{x}}_{\text{WLS}} = \hat{\mathbf{x}}_{\text{LS}} = \text{argmin}_{\mathbf{x}} \mathscr{L}_{\text{LS}}(\mathbf{x}) = \text{argmin}_{\mathbf{x}} \mathscr{L}_{\text{WLS}}(\mathbf{x})$$

Minimizing the *Weighted* Least Squares Criterion

Expanding our new criterion,

$$\begin{aligned}\mathscr{L}_{\mathrm{WLS}}(\mathbf{x}) &= \mathbf{e}^T\mathbf{R}^{-1}\mathbf{e}\\ &= (\mathbf{y}-\mathbf{H}\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x})\end{aligned}$$

We can minimize it as before, but accounting for the new weighting term:

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \mathscr{L}(\mathbf{x}) \longrightarrow \left.\frac{\partial \mathscr{L}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}} = \mathbf{0} = -\mathbf{y}^T\mathbf{R}^{-1}\mathbf{H} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$$

$$\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{x}}_{\mathrm{WLS}} = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}$$

The weighted normal equations!

Method of *Weighted* Least Squares

The weighted normal equations

$$\boxed{\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}}$$

Let's adapt our example:

	Resistance Measurements (Ohms)	
#	Multimeter A ($\sigma = 20$ Ohms)	Multimeter B ($\sigma = 2$ Ohms)
1	1068	
2	988	
3		1002
4		996

Method of *Weighted* Least Squares

Once we define the relevant quantities, we plug-and-chug to get:

$$\mathbf{H} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 1068 \\ 988 \\ 1002 \\ 996 \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{bmatrix} = \begin{bmatrix} 400 & & & \\ & 400 & & \\ & & 4 & \\ & & & 4 \end{bmatrix}$$

$$\begin{aligned}
\hat{x}_{\mathrm{WLS}} &= (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \\
&= \left([1111] \begin{bmatrix} 400 & & & \\ & 400 & & \\ & & 4 & \\ & & & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right)^{-1} [1111] \begin{bmatrix} 400 & & & \\ & 400 & & \\ & & 4 & \\ & & & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1068 \\ 988 \\ 1002 \\ 996 \end{bmatrix} \\
&= \frac{1}{1/400 + 1/400 + 1/4 + 1/4} \left(\frac{1068}{400} + \frac{988}{400} + \frac{1002}{4} + \frac{996}{4} \right) \\
&= 999.3 \text{ Ohms}
\end{aligned}$$

Ordinary versus Weighted Least Squares

	Least Squares	**Weighted Least Squares**
Loss / Criterion	$\mathscr{L}_{\mathrm{LS}}(\mathbf{x}) = \mathbf{e}^T\mathbf{e}$	$\mathscr{L}_{\mathrm{WLS}}(\mathbf{x}) = \mathbf{e}^T\mathbf{R}^{-1}\mathbf{e}$
Solution	$\hat{\mathbf{x}}_{\mathrm{LS}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$	$\hat{\mathbf{x}}_{\mathrm{WLS}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}$
Limitations	$m \geq n$	$m \geq n$ $\sigma_i^2 > 0$

Accurate noise modelling is crucial to utilize various sensors effectively!

Summary | Weighted Least Squares

- Measurements can come from sensors that have different noisy characteristics
- Weighted least squares lets us weight each measurement according to noise *variance*