Weighted Least Squares

Measurements prone to error.

Recall: $A\vec{x} = \vec{b}$

Choose x s.t $\|b - Ax\|^2$ is minimized

If $b \in$ col sp(A), then $\|b - Ax\|^2 = 0$

Error: $\|b - Ax\|^2 = 0$ if $b \in$ Col Sp(A).

$\neq 0$ if $b \notin$ col. sp(A).

$\Rightarrow$ $\|b - A\hat{x}\|^2$ $A\hat{x}$: Proj. of b onto the col·sp(A).

$\Rightarrow$ We will not solve $A\vec{x} = \vec{b}$. Instead, we solve the normal eqn

$$\underline{A^T A}\,\hat{x} = \underline{A^T} b.$$

If $\vec{b}$ has measurement errors $\Rightarrow$ Measurement errors in $\vec{b}$ are indep r.vs. with mean 0 & Variance = 1. then minimizing $\|b - A\hat{x}\|^2$ makes Sense.

$\underbrace{\text{r.vs. with mean 0 \& Variance = 1}}_{\text{Gaussian / Normal.}}$

Suppose if the errors are not indep & are of different Variances then $\|b - A\hat{x}\|^2$ minimizn does not make sense.

Measuring resistance values.

Measurement	1	2	3	4	. . .
R. Resistance Val. (Ω)	102	98	104	95	

What is the true value of R?

True Resistance   Noise.

$$
\left.
\begin{aligned}
y_1 &= x\,\Omega + \nu_1 \\
y_2 &: x + \nu_2 \\
y_3 &: x + \nu_3 \\
y_4 &: x + \nu_4.
\end{aligned}
\right\}
$$

Cost fn: Minimize : $(y_1 - x)^2 + (y_2 - x)^2 + \dots + (y_4 - x)^2.$

$$e^2 = \sum_{i=1}^{4} \left(\vec{y}_i - \vec{x}\right)^2 = \|\vec{y} - \vec{x}\|^2 = (y - x)^T (y - x).$$

* Measure the resistance value across 2 multimeter

M1: Variance $10\,\Omega$ M2 Variance $1\,\Omega$.

Since $\sigma_B^2 = 1$ & $\sigma_A^2 = 10$, Multimeter B is more reliable than A.

$\Rightarrow$ Give more importance to value measured by B & less to the value measured by B.

Incorporating this idea. we get the error that has to be minimized is given as follows.

$$\text{Minimize} \quad \sum_{i=1}^{n} \left(\frac{b - Ax}{\sigma_i^2}\right)_i^2$$

A function that needs to be optimized is called the objective fn / cost fn.

Each b_i has zero mean & Variance $= \sigma_i^2$. $\rightarrow$ Indep Measurement

$$z = \frac{X - \mu}{\sigma}$$

Mean & Var(z)

0   1

$$\therefore \quad a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

$m > n$

& no soln

Divide each eqn by σ_i, we have the variance of $\frac{b_i}{\sigma_i} = 1$.

$$V = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2 \end{bmatrix}_{m \times m}$$

$$V^{1/2} =$$

$$A\vec{x} = \vec{b}$$

Dividing both sides by σ_i we get

$$\Rightarrow \quad V^{-1/2} A\vec{x} = V^{-1/2}\vec{b}$$

where $V^{-1/2} = \begin{bmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_m} \end{bmatrix}$

$$\boxed{V^{-1/2} A \hat{x} = V^{-1/2} b}$$

$$\underbrace{V^{-1/2} A}_{} x = V^{-1/2} b.$$

$$Mx = b.$$

$$M^T M x = M^T b.$$

Apply ordinary least sq. here.

$$(V^{-1/2} A)^T V^{-1/2} A x = (V^{-1/2} A)^T V^{-1/2} b.$$

$$\Rightarrow \quad A^T V^{-\frac{1}{2}} . V^{-1/2} A x = A^T V^{-1/2} V^{-1/2} b$$

$$= \quad \rightarrow \boxed{A^T V^{-1} A x = A^T V^{-1} b.}$$

Expression for WLS.

$$V = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2 \end{bmatrix} \Rightarrow V^{-1} = \begin{bmatrix} 1/\sigma_1^2 & \cdots & 0 \\ 0 & & \frac{1}{\sigma_m^2} \end{bmatrix}$$

Since V^{-1} has elements $1/\sigma_i^2$, more reliable equations (with smaller variance) have larger weights.

Recall : If $b \notin$ col. sp(A), we can only estimate $\vec{b}$ by taking the proj of b onto col. sp(A).

$$A\vec{x} = b \quad \Rightarrow \quad A\underline{\hat{x}} = P\vec{b}$$

Find the variance of $\hat{x}$ in order to ascertain the reliability of the whole expt.

$$b = A\hat{x}.$$

If b is zero mean $\Rightarrow$ $\hat{x}$ has also zero mean.

$$y = 5x.$$

$E[x] = 0 \quad E(y) = ?$

CovVar Matrix : $(\hat{x}) = \underbrace{E\left[(\hat{x}-x)(\hat{x}-x)^T\right]}$.

$$\text{Cov} : E\left[(X-\bar{X})(X-\bar{X})^T\right]$$

$$\Downarrow$$

$$\boxed{\text{Cov}(\hat{x}) = (A^T V^{-1} A)^{-1}}$$

$$y = (a_1 x_1 + a_2 x_2) \qquad \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

$$\mathrm{Var}(Y) = a_1^2 \, \mathrm{Var}(X_1) + a_2^2 \, \mathrm{Var}(X_2) + 2 a_1 a_2 \, \mathrm{cor}(X_1, x_2).$$

$$= \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{cor}(X_1, X_2) \\ \mathrm{cov}(X_1, X_2) & \mathrm{Var}(X_2) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

$$y = (c_1 x_1 + c_2 x_2) = (b_{11} a_1 + b_{12} a_2) x_1 + (b_{21} a_1 + b_{22} a_2) x_2.$$

$$\mathrm{Var}(Y) = \begin{pmatrix} c_1 & c_2 \end{pmatrix} V \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

$$= \begin{pmatrix} b_{11} a_1 + b_{12} a_2 & b_{21} a_1 + b_{22} a_2 \end{pmatrix} V \begin{pmatrix} b_{11} a_1 + b_{12} a_2 \\ b_{21} a_2 + b_{22} a_2 \end{pmatrix}$$

$$= \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} V \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$\boxed{\vec{a}^T B^T V B \vec{a}.}$$

Ex: Suppose a doctor measures your bp 3 times and gets the following numbers.

$x = b_1, \quad x = b_2 \quad x = b_3$

$$1x = b_1$$
$$1x = b_2$$
$$1x = b_3$$

$A = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{x} = x. \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$

$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} \quad \sigma_1^2 = \frac{1}{9} \quad \sigma_2^2 = \frac{1}{4} \quad \& \ \sigma_3^2 = 1$

$V = \begin{pmatrix} 1/9 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad V^{-1} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$V^{-1/2} A \hat{x} = V^{-1/2} b$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{x} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$= \begin{aligned} 3\hat{x} &= 3b_1 \\ 2\hat{x} &= 2b_2 \\ 1\hat{x} &= 1\, b_3. \end{aligned}$$

$$A^T V^{-1} A x = A^T V^{-1} b$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 9 & & \\ & 4 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{x} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}9 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}\hat{x} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}9 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$$

$$= \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}9x\\4x\\x\end{bmatrix} = \begin{bmatrix}9b_1 + 4b_2 + b_3\end{bmatrix}$$

$$= 14\hat{x} = 9b_1 + 4b_2 + b_3$$

$$\boxed{\hat{x} = \frac{9b_1 + 4b_2 + b_3}{14}}$$

$\Rightarrow$ Weighted average of b_1, b_2, b_3

Ex: 2 $\quad x = b_1$ $\quad$ Variances σ_1^2 & σ_2^2.

$x = b_2$.

Find the best estimate $\hat{x}$ based on b_1 & b_2.

$$A\hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \hat{x} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad V = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$V^{-1/2} A\hat{x} = V^{-1/2} b$$

$$= \begin{pmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \hat{x} = \begin{pmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Normal eqn: $A^T V^{-1} \hat{x} = A^T V^{-1} b$.

$$= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} x = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\Rightarrow \frac{x}{\sigma_1^2} + \frac{x}{\sigma_2^2} = \frac{b_1}{\sigma_1^2} + \frac{b_2}{\sigma_2^2}$$

$$\boxed{\hat{x} = \frac{\frac{b_1}{\sigma_1^2} + \frac{b_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}}$$

WLSE.

If $\sigma_2 = 0 \Rightarrow \hat{x} = \dfrac{\sigma_2^2 b_1 + \sigma_1^2 b_2}{\sigma_1^2 + \sigma_2^2}$

$$= \frac{\sigma_1^2 b_2}{\sigma_1^2} = b_2.$$

If $\sigma_2 = \infty \Rightarrow \hat{x} = b_1$

RWLS / Kalman / GMM.

Calculus → Unit 4 : Diff, Gradient, Hessian

Derivatives of Matrices

Uncont.

Unit 6. LS, GDA, Optimizⁿ → Constraints

Lagrangian