Derivation of the Optimal LPQ Estimator in the Intercept-Only Model

1. Model Setup

Consider the intercept-only model

$$y_i = \mu + j_i, \quad i = 1, \dots, n,$$

with i.i.d. errors j_i satisfying

$$E[j_i] = 0, \quad \mathrm{Var}(j_i) = \sigma^2, \quad E[j_i^3] = m_3, \quad E[j_i^4] = m_4.$$

The usual unbiased estimator is $\bar{y} = \frac{1}{n}\sum_i y_i$ with $\mathrm{Var}(\bar{y}) = \sigma^2/n$. We seek an improved equivariant unbiased estimator of the form

$$\tilde{\mu} \;=\; \frac{1}{n}\sum_{i=1}^{n} y_i \;+\; B, \quad B = -\frac{1}{\sigma}\big(y^\top H y\big) + \sigma\,\mathrm{tr}(H),$$

where H is an $n \times n$ symmetric matrix to be chosen.

Equivariance/unbiasedness constraints: - Unbiasedness: $E[\tilde{\mu}] = \mu$. Since $E[y_i] = \mu$, the linear part $\frac{1}{n}\sum y_i$ is unbiased. For B,

$$E[B] = -\frac{1}{\sigma}E[y^\top H y] + \sigma\,\mathrm{tr}(H).$$

Write $y = \mu\mathbf{1} + j$. Then

$$y^\top H y = (\mu\mathbf{1} + j)^\top H(\mu\mathbf{1} + j) = \mu^2\,\mathbf{1}^\top H\mathbf{1} \;+\; 2\mu\,\mathbf{1}^\top H j \;+\; j^\top H j.$$

To ensure $E[B]$ does not depend on μ (so unbiased for all μ), require

$$H\,\mathbf{1} = 0 \quad\Longrightarrow\quad \mathbf{1}^\top H\mathbf{1} = 0,\; \mathbf{1}^\top H j = 0,$$

so $y^\top H y = j^\top H j$. Then

$$E[y^\top H y] = E[j^\top H j] = \sum_i H_{ii}E[j_i^2] + \sum_{i\neq j} H_{ij}E[j_i j_j] = \sigma^2\,\mathrm{tr}(H).$$

Hence $E[B] = -\frac{1}{\sigma}(\sigma^2\mathrm{tr}(H)) + \sigma\mathrm{tr}(H) = 0$. Thus unbiasedness holds whenever $H\mathbf{1} = 0$.

By symmetry of the i.i.d. setting, take H with

$$H_{ii} = a \quad (\forall i), \;\; H_{ij} = b \quad (\forall i \neq j), \;\; a + (n-1)b = 0 \;\Longrightarrow\; b = -\frac{a}{n-1}.$$

2. Variance decomposition

Write
$$\tilde{\mu} = \mu + A + B, \quad A = \frac{1}{n}\sum_{i=1}^{n} j_i, \quad B = -\frac{1}{\sigma}(j^\top H j) + \sigma\,\mathrm{tr}(H).$$
Since $E[A] = 0$, $E[B] = 0$,
$$\mathrm{Var}(\tilde{\mu}) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\,\mathrm{Cov}(A, B).$$

2.1. Var(A)

$$\mathrm{Var}(A) = \mathrm{Var}\Big(\frac{1}{n}\sum_i j_i\Big) = \frac{\sigma^2}{n}.$$

2.2. Cov(A, B)

$$\mathrm{Cov}\Big(A,\ -\frac{1}{\sigma} j^\top H j\Big) = -\frac{1}{n\sigma}\,\mathrm{Cov}\Big(\sum_{i=1}^{n} j_i,\ \sum_{k,\ell} H_{k\ell}\, j_k j_\ell\Big).$$
Only diagonal terms $k = \ell$ contribute under independence and zero mean:
$$\mathrm{Cov}\Big(\sum_i j_i,\ j^\top H j\Big) = \sum_{k=1}^{n} H_{kk} E[j_k^3] = (na)\, m_3.$$
Thus
$$\mathrm{Cov}(A, B) = -\frac{1}{n\sigma}\,(nam_3) = -\frac{a\,m_3}{\sigma}, \quad 2\,\mathrm{Cov}(A, B) = -\,\frac{2a\,m_3}{\sigma}.$$

2.3. Var(B)

We need $\mathrm{Var}(j^\top H j)$. As before,
$$j^\top H j = a\sum_{i=1}^{n} j_i^2 \ +\ b\sum_{i\neq j} j_i j_j, \quad b = -\frac{a}{n-1}.$$
- $\mathrm{Var}\big(a\sum_i j_i^2\big) = a^2\sum_i \mathrm{Var}(j_i^2) = a^2 n\big(m_4 - \sigma^4\big)$. - $\mathrm{Var}\big(b\sum_{i\neq j} j_i j_j\big) = b^2\, n(n-1)\sigma^4$. No covariance between the two sums. Hence
$$\mathrm{Var}(j^\top H j) = a^2 n(m_4 - \sigma^4) + b^2 n(n-1)\sigma^4.$$
Thus
$$\mathrm{Var}(B) = \mathrm{Var}\Big(-\frac{1}{\sigma} j^\top H j\Big) = \frac{1}{\sigma^2}\,\mathrm{Var}(j^\top H j) = \frac{a^2 n(m_4 - \sigma^4) + b^2 n(n-1)\sigma^4}{\sigma^2}.$$
Substitute $b = -a/(n-1)$:
$$b^2 n(n-1)\sigma^4 = \frac{na^2\sigma^4}{n-1},$$
so
$$\mathrm{Var}(B) = \frac{a^2}{\sigma^2}\Big[n(m_4 - \sigma^4) + \frac{n\sigma^4}{n-1}\Big].$$

2.4. Total variance and minimization

Combine:

$$\mathrm{Var}(\tilde{\mu}) = \frac{\sigma^2}{n} \;+\; \frac{a^2}{\sigma^2}\Big[n(m_4-\sigma^4) + \tfrac{n\sigma^4}{n-1}\Big] \;-\; \frac{2a\,m_3}{\sigma}.$$

Let

$$C = n(m_4-\sigma^4) + \frac{n\sigma^4}{n-1}.$$

Then

$$\mathrm{Var}(\tilde{\mu}) = \frac{\sigma^2}{n} + \frac{a^2}{\sigma^2}C - \frac{2a\,m_3}{\sigma}.$$

Differentiate w.r.t. a:

$$\frac{d}{da}\Big(\frac{a^2}{\sigma^2}C - \frac{2a\,m_3}{\sigma}\Big) = 2\frac{a}{\sigma^2}C - 2\frac{m_3}{\sigma} = 0 \implies a = \frac{m_3\,\sigma}{C}.$$

Thus the optimal H has

$$H_{ii} = a = \frac{m_3\,\sigma}{C}, \quad H_{ij} = -\frac{a}{n-1} = -\frac{m_3\,\sigma}{(n-1)\,C},\ i \neq j.$$

Plugging back:

$$\frac{a^2}{\sigma^2}C = \frac{m_3^2\sigma^2}{C^2}\cdot\frac{C}{\sigma^2} = \frac{m_3^2}{C}, \quad -\frac{2am_3}{\sigma} = -\frac{2m_3^2}{C}.$$

Hence

$$\mathrm{Var}(\tilde{\mu}_{\mathrm{opt}}) = \frac{\sigma^2}{n} + \frac{m_3^2}{C} - \frac{2m_3^2}{C} = \frac{\sigma^2}{n} - \frac{m_3^2}{C}.$$

Recalling $C = n(m_4-\sigma^4) + \frac{n\sigma^4}{n-1}$, the final general formula is

$$\boxed{\mathrm{Var}(\tilde{\mu}_{\mathrm{opt}}) = \frac{\sigma^2}{n} \;-\; \frac{m_3^2}{n\big(m_4-\sigma^4\big) + \frac{n\sigma^4}{n-1}}.}$$

3. Remarks

- If $m_3 = 0$ (symmetric errors), then $a = 0$, so $H = 0$ and $\tilde{\mu} = \bar{y}$. - When $m_3 \neq 0$, one obtains variance reduction below σ^2/n, the amount governed by m_3^2 and m_4. - The matrix H remains exchangeable ($H_{ii} = a$, $H_{ij} = -a/(n-1)$) ensuring $H\mathbf{1} = 0$.