Admissibility

Consider estimators \( \mathbf{m} \) and \( \mathbf{m}^* \) of \( \mu \):

- The sum-of-squared-errors loss function is
  \[
  L(\mu, \mathbf{m}) = (\mathbf{m} - \mu)'(\mathbf{m} - \mu) = \|\mathbf{m} - \mu\|^2.
  \]

- The risk function \( R(\mu, \mathbf{m}) \) is the expected loss:
  \[
  R(\mu, \mathbf{m}) = \mathbb{E}_\mu[L(\mu, \mathbf{m})]
  \]

- \( \mathbf{m}^* \) is as good as \( \mathbf{m} \) if
  \[
  \forall \mu, R(\mu, \mathbf{m}^*) \leq R(\mu, \mathbf{m})
  \]

- \( \mathbf{m}^* \) is better than \( \mathbf{m} \) if it is as good, and
  \[
  \exists \mu, R(\mu, \mathbf{m}^*) < R(\mu, \mathbf{m})
  \]

- \( \mathbf{m} \) is admissible if there is no \( \mathbf{m}^* \) that is better than \( \mathbf{m} \).
Bayes estimation

- If \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \) is a random sample from \( \mathcal{N}_p(\mu, \Sigma) \), and \( \mu \) has the prior distribution \( \mathcal{N}_p(\nu, \Phi) \), then the posterior distribution of \( \mu \) is multivariate normal with mean

\[
\Phi \left( \Phi + \frac{1}{N} \Sigma \right)^{-1} \bar{x} + \frac{1}{N} \Sigma \left( \Phi + \frac{1}{N} \Sigma \right)^{-1} \nu
\]

and covariance matrix

\[
\Phi - \Phi \left( \Phi + \frac{1}{N} \Sigma \right)^{-1} \Phi
\]

- The posterior mean of \( \mu \) is “a kind of weighted average of \( \bar{x} \) and \( \nu \).”
The posterior mean minimizes the expected risk

\[ r(\nu, \Phi, m) = \mathbb{E}_{\nu, \Phi}[R(\mu, m)] \]

and is called the \textit{Bayes estimator}.

- Any Bayes estimator is admissible.
- Any admissible estimator is a Bayes estimator, or the limit of Bayes estimators.

An estimator \( m \) is \textit{minimax} if

\[ \sup_{\mu} R(\mu, m) = \inf_{m^*} \sup_{\mu} R(\mu, m^*) \]

\( \bar{x} \) is minimax.
Improved Estimation of the Mean Vector

As an estimator of \( \mu \) in a random sample of \( N \) from \( N_p(\mu, \Sigma) \), \( \bar{x} \) is:

- the maximum likelihood estimate;
- minimum variance unbiased;
- equivariant: \( A\bar{x} + b = A\bar{x} + b \).

But it is not admissible with respect to sum-of-squared-errors loss, when \( \Sigma \propto I_p \) and \( p \geq 3 \).
The James-Stein estimator:

- Take $\Sigma = N I_p$, so that $\bar{X} \sim N_p(\mu, I_p)$.
- Then $E[L(\mu, \bar{X})] = p$.
- Now for any fixed $\nu$, let

$$m(\bar{x}) = \left(1 - \frac{p - 2}{||\bar{x} - \nu||^2}\right)(\bar{x} - \nu) + \nu$$

- This estimator shrinks $\bar{x}$ toward the arbitrary $\nu$.
- If $p \geq 3$, $m(\bar{x})$ is better than $\bar{x}$, proving that $\bar{x}$ is not admissible.
Note that if \( ||\bar{x} - \nu||^2 = p - 2 \), \( m(\bar{x}) = \nu \); that is, the shrinkage has gone all the way from \( \bar{x} \) to \( \nu \).

If \( ||\bar{x} - \nu||^2 < p - 2 \), the “shrinkage” goes beyond \( \nu \).

A more sensible estimator is

\[
m^+(\bar{x}) = \left( 1 - \frac{p - 2}{||\bar{x} - \nu||^2} \right)^+ (\bar{x} - \nu) + \nu
\]

where \((x)^+\) denotes the \textit{positive part} of \( x \):

\[
(x)^+ = \begin{cases} 
    x & x \geq 0 \\
    0 & x < 0
\end{cases}
\]

- \( m^+(\bar{x}) \) is better than \( m(\bar{x}) \).
- \( m^+(\bar{x}) \) is minimax.
- But \( m^+(\bar{x}) \) is also not admissible.