# MAXIMUM LIKELIHOOD ESTIMATION OF AN INCREASING NONNEGATIVE SIGNAL

(10 points possible)

We wish to estimate a scalar signal $x(t)$, for $t = 1, 2, \ldots, N$, which is known to be nonnegative and monotonically nondecreasing:

$$0 \leq x(1) \leq x(2) \leq \cdots \leq x(N).$$

This occurs in many practical problems. For example, $x(t)$ might be a measure of wear or deterioration, that can only get worse, or stay the same, as time $t$ increases. We are also given that $x(t) = 0$ for $t \leq 0$.

We are given a noise-corrupted moving average of $x$, given by

$$y(t) = \sum_{\tau=1}^{k} h(\tau)x(t-\tau) + v(t), \quad t = 2, \ldots, N+1,$$

where $v(t)$ are independent $\mathcal{N}(0,1)$ random variables.

Formulate the problem of finding the maximum likelihood estimate of $x$, given $y$, taking into account the prior assumption that $x$ is nonnegative and monotonically nondecreasing, as a convex optimization problem. Now solve a specific instance of the problem, with problem data (i.e., $N$, $k$, $h$, and $y$) given in the file `ml_estim_incr_signal_data_norng.m`. (This file contains the true signal `xtrue`, which of course you cannot use in creating your estimate.) Find the maximum likelihood estimate $\hat{x}_{\mathrm{ml}}$, and plot it, along with the true signal. Also find and plot the maximum likelihood estimate $\hat{x}_{\mathrm{ml,free}}$ *not taking into account the signal nonnegativity and monotonicity*.

*Hint.* The function `conv` (convolution) is overloaded to work with CVX.

Which of the following statements most accurately describe the plot?

- It turns out that $\hat{x}_{\mathrm{ml}} = \hat{x}_{\mathrm{ml,free}}$, which implies that the nonnegativity and monotonicity constraints were, in fact, unnecessary.
- It turns out that $\hat{x}_{\mathrm{ml,free}}$ is not monotonous but nonnegative, which implies that the nonnegativity constraint was, in fact, unnecessary.
- The maximum deviation of $\hat{x}_{\mathrm{ml}}$ from the true signal, $x$, is smaller than that of $x_{\mathrm{ml,free}}$.
- The maximum deviation of $\hat{x}_{\mathrm{ml}}$ from the true signal, $x$, is larger than that of $x_{\mathrm{ml,free}}$. This, however, does not mean $x_{\mathrm{ml,free}}$ is a better estimate because $x_{\mathrm{ml,free}}$ does not satisfy the constraints and therefore is not a sensible estimate.

Show Answer

*You have used 0 of 2 submissions*

Help

## WORST-CASE PROBABILITY OF LOSS (20 points possible)

Two investments are made, with random returns $R_1$ and $R_2$. The total return for the two investments is $R_1 + R_2$, and the probability of a loss (including breaking even, i.e., $R_1 + R_2 = 0$) is $p^{\text{loss}} = \textbf{prob}(R_1 + R_2 \leq 0)$. The goal is to find the worst-case (i.e., maximum possible) value of $p^{\text{loss}}$, consistent with the following information. Both $R_1$ and $R_2$ have Gaussian marginal distributions, with known means $\mu_1$ and $\mu_2$ and known standard deviations $\sigma_1$ and $\sigma_2$. In addition, it is known that $R_1$ and $R_2$ are correlated with correlation coefficient $\rho$, i.e.,

$$\mathbf{E}(R_1 - \mu_1)(R_2 - \mu_2) = \rho\sigma_1\sigma_2.$$

Your job is to find the worst-case $p^{\text{loss}}$ over any joint distribution of $R_1$ and $R_2$ consistent with the given marginals and correlation coefficient.

We will consider the specific case with data

$$\mu_1 = 8, \quad \mu_2 = 20, \quad \sigma_1 = 6, \quad \sigma_2 = 17.5, \quad \rho = -0.25.$$

We can compare the results to the case when $R_1$ and $R_2$ are jointly Gaussian. In this case we have

$$R_1 + R_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2),$$

which for the data given above gives $p^{\text{loss}} = 0.050$. Your job is to see how much larger $p^{\text{loss}}$ can possibly be.

This is an infinite-dimensional optimization problem, since you must maximize $p^{\text{loss}}$ over an infinite-dimensional set of joint distributions. To (approximately) solve it, we discretize the values that $R_1$ and $R_2$ can take on, to $n = 100$ values $r_1, \ldots, r_n$, uniformly spaced from $r_1 = -30$ to $r_n = +70$. We use the discretized marginals $p^{(1)}$ and $p^{(2)}$ for $R_1$ and $R_2$, given by

$$p_i^{(k)} = \textbf{prob}(R_k = r_i) = \frac{\exp\left(-(r_i - \mu_k)^2/(2\sigma_k^2)\right)}{\sum_{j=1}^n \exp\left(-(r_j - \mu_k)^2/(2\sigma_k^2)\right)},$$

for $k = 1, 2$, $i = 1, \ldots, n$.

Formulate the (discretized) problem as a convex optimization problem, and solve it. What is the maximum value of $p^{\text{loss}}$?

Plot the joint distribution that yields the maximum value of $p^{\text{loss}}$ using the Matlab commands `mesh` and `contour`. Which of the following statements most accurately describe the plot of the (worst-case) joint distribution?

- The distribution has mass on the line where R1 + R2 = 0 (i.e., break even, which counts as a loss for us). Then it has extra mass on another region in the plane, which is needed to make the marginals the given Gaussians, as well as to meet the constraint on the correlation coefficient.
- The distribution looks like the joint Gaussian distribution. However, with more careful inspection, we see that the tails are heavier than the joint Gaussian (i.e., the density decays to $0$ at a rate slower than a Gaussian as you move away from the center).
- The distribution has several disjoint sharp peaks scattered in a seemingly random fashion. This suggests that

$R_1$ and $R_2$ is a Gaussian mixture.

Show Answer *You have used 0 of 2 submissions*