EE364a Review Session 6

topics:

- ML prediction with highly quantized measurements
- two-way partitioning

Estimation with quantized measurements

given:

- a signal matrix $A \in \mathbf{R}^{m \times n}$
- measurements $y = \phi(Ax + v)$, where $v \sim \mathcal{N}(0, \sigma^2 I)$ and

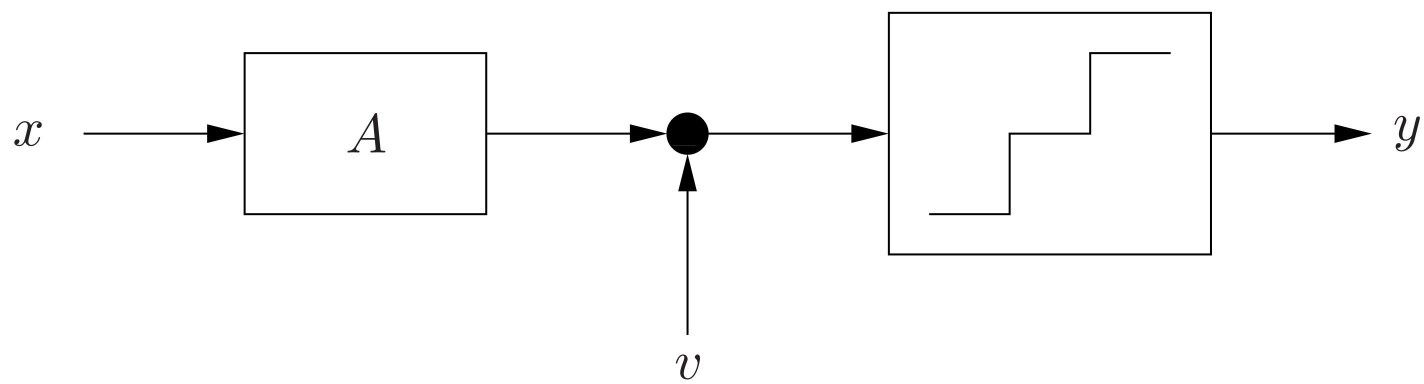
$$\phi_i : \mathbf{R} \to \{1, \ldots, K\}$$
$$\phi_i^{-1}(k) = (t_k, t_{k+1}]$$

- quantization levels

$$-\infty = t_1 < t_2 < t_3 < \cdots < t_K < t_{K+1} = \infty$$

compute $\hat{x}$, the maximum likelihood estimate of x, given y

Estimation with quantized measurements



how would you find $\hat{x}$

- with no noise or quantization ($v = 0$ and $\phi(z) = z$)?
- with noise, but not quantization ($\phi(z) = z$)?
- with no noise, but quantization ($v = 0$)?

Likelihood and log-likelihood

- likelihood:

$$p(y|x) = \prod_{i=1}^{m} \left(\Phi\left(\frac{t_{y_i+1} - (Ax)_i}{\sigma} \right) - \Phi\left(\frac{t_{y_i} - (Ax)_i}{\sigma} \right) \right)$$

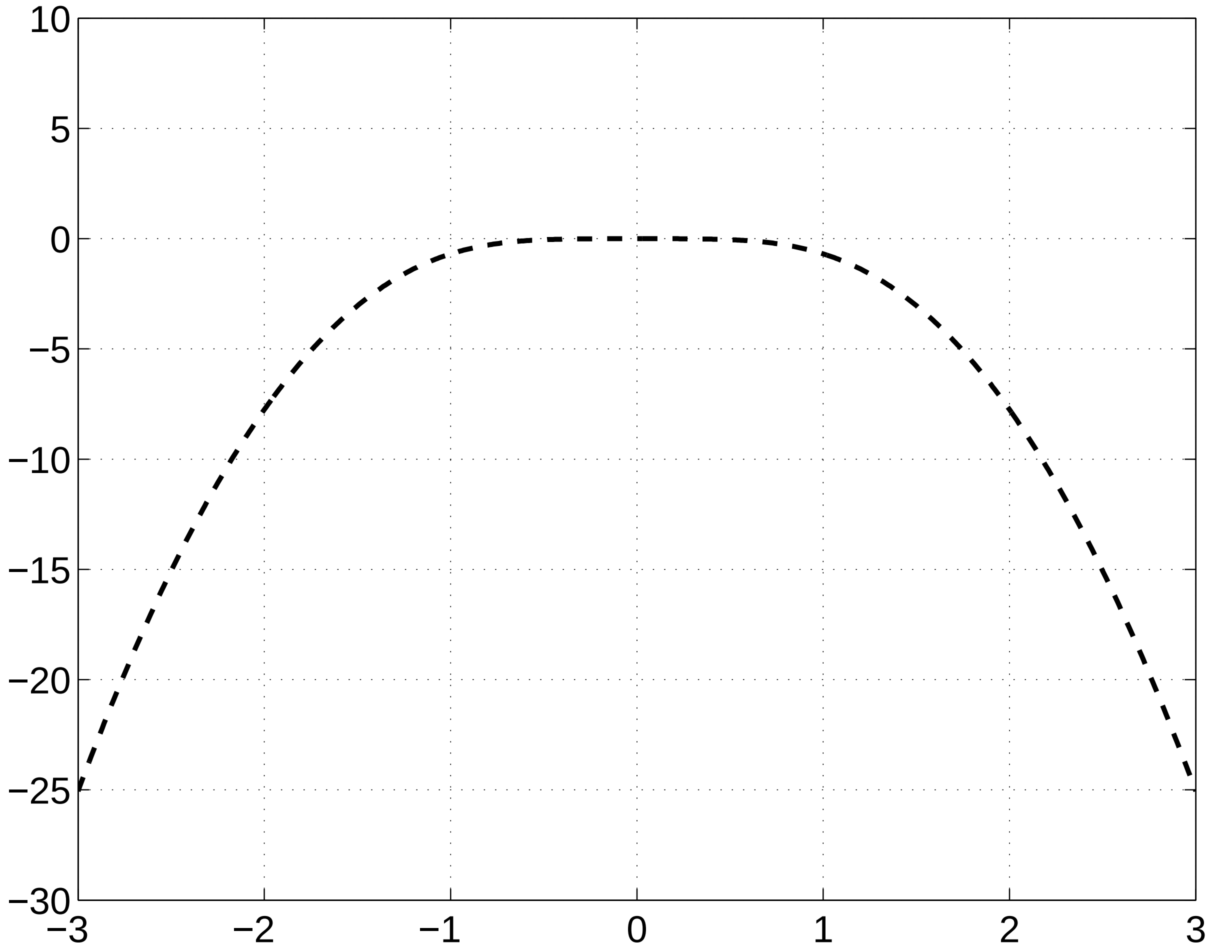
- log-likelihood:

$$l_y(x) = \sum_{i=1}^{m} \log \left(\Phi\left(\frac{t_{y_i+1} - (Ax)_i}{\sigma} \right) - \Phi\left(\frac{t_{y_i} - (Ax)_i}{\sigma} \right) \right)$$

where Φ is the cdf of the standard normal distribution

- $l_y(x)$ is concave, twice differentiable

Interval log-normal cdf



plot of $f(x) = \log(\Phi((x+1)/\sigma) - \Phi((x-1)/\sigma))$, for $\sigma = 0.3$

ML estimation

$$\text{maximize} \quad l_y(x)$$

- convex, unconstrained optimization problem
- can be efficiently solved using Newton's method (next topic)

extensions:

- MAP, with prior distribution on x
- prior constraints on x

Numerical example

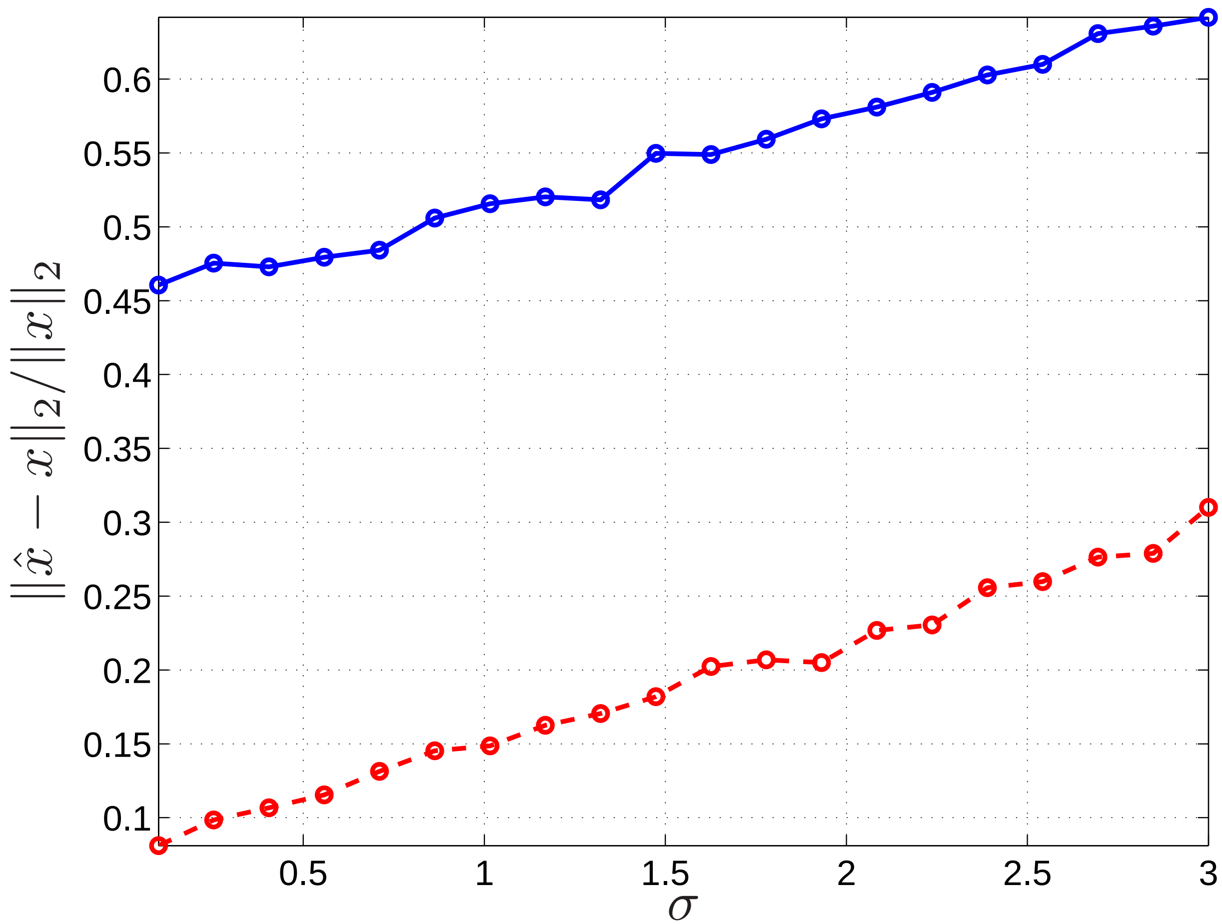
problem instance:

- $n = 10$ variables, $m = 200$ measurements
- thresholds $-\infty, -1, +1, \infty$ (3 intervals ≈ 1.6 bits per measurement)
- $A_{ij} \sim \mathcal{N}(0, 1)$

simulation:

- vary σ from 0.1 to 3
- generate 100 values of x, y, with $x \sim \mathcal{N}(0, I)$
- compute $\hat{x}$
- evaluate relative estimation error $\|\hat{x} - x\|_2 / \|x\|_2$

Results



dashed: ML; solid: least-square, taking $y_i \in \{-2, 0, +2\}$

Two-way partitioning

- n vertices, labeled $\{1, \ldots, n\}$
- we are given a set of symmetric weights on pairs of vertices, $w_{ij} = w_{ji}$
- find partition of vertices (Y, Z)
  ($i.e.$, $Y \cup Z = \{1, \ldots, n\}$, $Y \cap Z = \emptyset$)
  which maximizes total weight of cut,

$$J(Y, Z) = \sum_{i \in Y} \sum_{j \in Z} w_{ij}$$

- encode partition via $x \in \{-1, 1\}^n$; $x_i = -1$ means $x \in Y$
- $J(x) = \mathbf{1}^T W \mathbf{1} - x^T W x$

Two-way partitioning

can be cast as

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1 \end{array}$$

or equivalently

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}(WX) \\ \text{subject to} & X_{ii} = 1, \quad X \succeq 0 \\ & \mathbf{rank}(X) = 1 \end{array}$$

- a nonconvex combinatorial problem
- we will derive an SDP relaxation

SDP relaxation

by dropping the rank constraint, we get

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}(WX) \\ \text{subject to} & X_{ii} = 1, \quad X \succeq 0 \end{array}$$

randomized scheme:

- solve SDP for $X^\star$ (gives lower bound)
- sample $v \sim \mathcal{N}(0, X^\star)$
- set $x = \mathbf{sign}(v)$

Goemans & Williamson proved that this lower bound is on average at most 14% suboptimal for the MAX-CUT problem ($W_{ii} = 0$, $W_{ij} \geq 0$)

SDP relaxation via dual

Lagrangian of original problem:

$$\begin{aligned} L(x,\nu) &= x^TWx + \sum_i \nu_i(x_i^2 - 1) \\ &= \mathbf{tr}\left((W + \mathbf{diag}(\nu))xx^T\right) - \mathbf{1}^T\nu \end{aligned}$$

dual function:

$$g(\nu) = \begin{cases} -\mathbf{1}^T\nu, & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty, & \text{otherwise} \end{cases}$$

SDP relaxation via dual

dual problem:

$$
\begin{array}{ll}
\text{maximize} & -\mathbf{1}^T \nu \\
\text{subject to} & W + \mathbf{diag}(\nu) \succeq 0
\end{array}
$$

dual of dual:

$$
\begin{array}{ll}
\text{minimize} & \mathbf{tr}(WX) \\
\text{subject to} & X_{ii} = 1, \quad X \succeq 0
\end{array}
$$

same as dropping the rank constraint!