4. Machine Learning and Binaural Hearing

1. Binaural Features
2. Mapping Sounds onto Their Directions
3. Collecting Training Data
4. The Binaural Manifold
5. Localization with a Look-up Table
6. **Linear Regression**
7. Piecewise Linear Regression
8. Predicting the Direction of Speech
9. Principles of Sound Separation
10. Separation & Localization Method

Multivariate Linear Regression (I)

- Consider the regression problem already described and applied to the training set: $\mathcal{T} = \{(\boldsymbol{y}_1, \boldsymbol{x}_1), \ldots, (\boldsymbol{y}_n, \boldsymbol{x}_n), \ldots, (\boldsymbol{y}_N, \boldsymbol{y}_N)\}$:
  1. Estimate $\hat{f}$ from:

$$\boldsymbol{x}_1 = f(\boldsymbol{y}_1),$$
$$\vdots$$
$$\boldsymbol{x}_N = f(\boldsymbol{y}_N).$$

  2. Predict the direction of a **white-noise emitter**:

$$\hat{\boldsymbol{x}} = \hat{f}(\boldsymbol{w})$$

Multivariate Linear Regression (II)

- Estimate matrix $\mathbf{A} \in \mathbb{R}^{2\times D}$ and vector $\boldsymbol{b} \in \mathbb{R}^2$ (affine transformation):

$$
\begin{aligned}
\boldsymbol{x}_1 &= \mathbf{A}\boldsymbol{y}_1 + \boldsymbol{b}, \\
&\vdots \\
\boldsymbol{x}_n &= \mathbf{A}\boldsymbol{y}_n + \boldsymbol{b}, \\
&\vdots \\
\boldsymbol{x}_N &= \mathbf{A}\boldsymbol{y}_N + \boldsymbol{b}.
\end{aligned}
$$

Linear Regression Formulation

- These equations can be rearranged to yield a matrix-vector equation of the form:

$$\mathbf{Y}\boldsymbol{a} = \boldsymbol{X}$$

- with:
  1. $\mathbf{Y} \in \mathbb{R}^{(2N)\times(2D+2)}$ is a matrix containing the observed input training data (ILPD vectors).
  2. $\mathbf{a} \in \mathbb{R}^{2D+2}$ is a vector containing the unknown entries of $\mathbf{A}$ and $\boldsymbol{b}$.
  3. $\boldsymbol{X} \in \mathbb{R}^{2N}$ is a vector containing the observed output training data (sound directions).

Least-Square Solution

- Each input-output pair $(\boldsymbol{y}_n, \boldsymbol{x}_n)$ yields **2 linear constraints**.
- A solution is possible if $\mathbf{Y} \in \mathbb{R}^{(2N)\times(2D+2)}$ is invertible, or if $N \geq D + 1$.
- The dimension of the ILPD vectors: $D = 1536$,
- A solution exists only if $N \geq D + 1 = 1537$ training pairs are available:

$$\hat{\boldsymbol{a}} = (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top \boldsymbol{X}$$

Interchanging Input and Output

- Linear regression may also be estimated **the other way around**:
- Estimate matrix $\mathbf{A} \in \mathbb{R}^{D \times 2}$ and vector $\boldsymbol{b} \in \mathbb{R}^{D}$:

$$\begin{aligned} \boldsymbol{y}_1 &= \mathbf{A}\boldsymbol{x}_1 + \boldsymbol{b}, \\ &\vdots \\ \boldsymbol{y}_n &= \mathbf{A}\boldsymbol{x}_n + \boldsymbol{b}, \\ &\vdots \\ \boldsymbol{y}_N &= \mathbf{A}\boldsymbol{x}_N + \boldsymbol{b}. \end{aligned}$$

Reversed Linear Regression

$$\boldsymbol{X}\boldsymbol{a} = \boldsymbol{Y}$$

with:

1. $\mathbf{X} \in \mathbb{R}^{(DN)\times(2D+D)}$ is a matrix containing the observed output training data (sound directions).
2. $\mathbf{a} \in \mathbb{R}^{2D+D}$ is a vector containing the unknown entries of $\mathbf{A}$ and $\boldsymbol{b}$.
3. $\boldsymbol{Y} \in \mathbb{R}^{DN}$ is a vector containing the observed input training data (ILPD vectors).

Reversed Least-Square Solution

- Each output-input pair $(\boldsymbol{x}_n, \boldsymbol{y}_n)$ yields D **linear constraints**.
- A solution is possible if $\mathbf{Y} \in \mathbb{R}^{(DN)\times(2D+D)}$ is invertible: $N \geq L + 1 = 3$.
- A minimum of $N = 3$ pairs are needed !
- With $N \gg 3$, the system is **over constrained**:

$$\hat{\boldsymbol{a}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \boldsymbol{Y}$$

Session Summary

- Multivariate linear regression
- Computational complexity
- Interchanging the input and the output
- Least-square solution