KHU-BME
Pattern Classification
Lecture 12

MAXIMUM A POSTERIORI MODE	We note in passing that a related class of estimators — *maximum a posteriori* or MAP estimators — find the value of $\boldsymbol{\theta}$ that maximizes $l(\boldsymbol{\theta})p(\boldsymbol{\theta})$. Thus a maximum likelihood estimator is a MAP estimator for the uniform or "flat" prior. As such, a MAP estimator finds the peak, or *mode* of a posterior density. The drawback of MAP estimators is that if we choose some arbitrary nonlinear transformation of the parameter space (e.g., an overall rotation), the density will change, and our MAP solution need no longer be appropriate (Sec. 3.5.2).

<u>Estimation of Random Parameters (General Methods)</u>

Parameters $\underline{\theta}$ are random variables, Choose $\underline{\hat{\theta}}$ by minimizaing the expected value of an error function or cost function between $\underline{\theta}$ and $\underline{\hat{\theta}}$

$\underline{z}$ = observed samples

Let $C(\underline{\theta},\underline{\hat{\theta}})$ be a cost function between $\underline{\theta}$ and $\underline{\hat{\theta}}$

The expected cost or risk of the estimate is

R=E{ $C(\underline{\theta},\underline{\hat{\theta}})$ }= $\iint C(\underline{\theta},\underline{\hat{\theta}}(z))p(z\,|\,\theta)dzd\theta$ *Note C(.) is not specified yet*

The estimate $\underline{\hat{\theta}}$ that minimizes R is called the Bayes estimate

$$R=\iint C(\underline{\theta},\underline{\hat{\theta}}(z))p(z\,|\,\theta)dzd\theta=\int p(z)[\int C(\theta,\hat{\theta})p(\theta\,|\,z)d\theta)dz$$ by the Bayes Rule

Bayes estimate is

- very general
- difficult to calculate for general cost functions

Now look at some a specific choice

R=∫ p($\underline{z}$) [∫ C($\underline{\theta}$, $\underline{\hat{\theta}}$($\underline{z}$)) p($\underline{\theta}$| $\underline{z}$) d$\underline{\theta}$] d$\underline{z}$

↓

$$1-\int_{\Delta\Gamma} p(\theta\,|\,z)d\theta$$

Minimize R => minimize [·] for each $\underline{z}$ => maximize p ($\underline{\hat{\theta}}$| $\underline{z}$)

MAP:

=>

Use Bayes theroem

$\ln p(\underline{\theta}|\,\underline{z}) = \ln p(\underline{z}\,|\underline{\theta}) + \ln p(\underline{\theta}) - \ln p(\underline{z})$

MAP estimate:

=>

Note If $p(\underline{\theta})$ is much more flatly distributed (vs. $\underline{\theta}$) than $p(\underline{z}\,|\underline{\theta})$, second term can be neglected => MAP = ML.

Estimation Summary

r = s + n

Measurement r ⇔ $\underline{z}$

Now estimate s ⇔ $\underline{\theta}$, parameter or signal

where n is noise

1. **ML**: $\hat{s}$ = a value that maximizes p{r|s} ⇔ $\hat{\underline{\theta}}$ = a value of $\underline{\theta}$ that maximizes $p\{\underline{z}|\underline{\theta}\}$

2. **MAP**: $\hat{s}$ = a value that maximizes p{s|r} ⇔ $\hat{\underline{\theta}}$ = a value of $\underline{\theta}$ that maximizes $p\{\underline{\theta}|\underline{z}\}$, a posteriori density.

The same concept can be applied to

$\mathbf{y} = \mathbf{x} + \mathbf{n} \Leftrightarrow \mathbf{r} = \mathbf{s} + \mathbf{n}$

where y is measurement (of signal or image), x is true value (of signal or image), and n is noise.

Bayesian Image Estimation (Restoration) Example

Before After

- G. K. Chantas, Bayesian Restoration Using a New Nonstaionary Edge-Preserving Image Prior, IEEE Trans. Image Processing, 15, 10, 2006