Contents

Chapter 1

Introduction

1.1 Curve Fitting

Problem 1

This can be solved by substituting the definition of:

$$y(x, \mathbf{w}) = \sum_{j=0}^{M} w_j x^j$$

into the error function and then taking the derivative.

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} \left(y(x, \mathbf{w}) - t_n \right)^2$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x^j - t_n \right)^2 \qquad \text{Substitute}$$

$$\frac{dE(\mathbf{w})}{dw_i} = \sum_{n=1}^{N} \left(\left(\sum_{j=0}^{M} w_j x^j - t_n \right) x^i \right) \qquad \text{Take the derivative}$$

$$0 = \sum_{n=1}^{N} \left(\left(\sum_{j=0}^{M} w_j x^j - t_n \right) x^i \right) \qquad \text{Set derivative to 0}$$

$$0 = \sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x^j x^i - t_n x^i \right) \qquad \text{Set derivative to 0}$$

$$\sum_{n=1}^{N} t_n x^i = \sum_{n=1}^{N} \sum_{j=0}^{M} w_j x^j x^i$$

$$\sum_{n=1}^{N} t_n x^i = \sum_{n=1}^{N} \sum_{j=0}^{M} w_j x^{i+j}$$

Problem 2

This is solved in almost the same way we just have one additional term for the regularization so:

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} (y(x, \mathbf{w}) - t_n)^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x^j - t_n \right)^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

$$\frac{dE(\mathbf{w})}{dw_i} = \sum_{n=1}^{N} \left(\left(\sum_{j=0}^{M} w_j x^j - t_n \right) x^i \right) + \lambda w_i$$

$$-\lambda w_i = \sum_{n=1}^{N} \left(\left(\sum_{j=0}^{M} w_j x^j - t_n \right) x^i \right) + \lambda w_i$$

$$-\lambda w_i = \sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x^j x^i - t_n x^i \right)$$

$$\sum_{n=1}^{N} t_n x^i - \lambda w_i = \sum_{n=1}^{N} \sum_{j=0}^{M} w_j x^j x^i$$

$$\sum_{n=1}^{N} t_n x^i - \lambda w_i = \sum_{n=1}^{N} \sum_{j=0}^{M} w_j x^{i+j}$$

1.2   Probability Theory

Problem 3 The Probability of Selecting an Apple can be decomposed as:

$$\mathcal{P}(apple) = \mathcal{P}(apple, red) + \mathcal{P}(apple, blue) + prob(apple, green)$$
$$= \mathcal{P}(apple|red)\mathcal{P}(red) + \mathcal{P}(apple|blue)\mathcal{P}(blue) + \mathcal{P}(apple|green)\mathcal{P}$$
$$= (.3)(.2) + (.5)(.2) + (.3)(.6)$$
$$= .34$$

The probability that observing an orange came from the green box

can be solved using Bayes rule:

$$P(green|orange) = \frac{P(orange|green)P(green)}{P(orange)}$$
$$= \frac{(.3)(.6)}{.66}$$
$$= .27$$

Problem 5

$$\text{Var}\left[X\right] = \mathbb{E}\left[(X - \mathbb{E}\left[X\right])^2\right]$$
$$= \mathbb{E}\left[X^2 - 2X\mathbb{E}\left[X\right] + \mathbb{E}\left[X\right]^2\right] \qquad \text{Distributive Law}$$
$$= \mathbb{E}\left[X^2\right] + \mathbb{E}\left[-2X\mathbb{E}\left[X\right]\right] + \mathbb{E}\left[\mathbb{E}\left[X\right]^2\right] \quad \text{Linearity of } \mathbb{E}\left[X\right]$$
$$= \mathbb{E}\left[X^2\right] + -2\mathbb{E}\left[X\mathbb{E}\left[X\right]\right] + \mathbb{E}\left[X\right]^2 \qquad \mathbb{E}\left[\alpha X\right] = \alpha \mathbb{E}\left[X\right]$$
$$= \mathbb{E}\left[X^2\right] + -2\mathbb{E}\left[X\right]^2 + \mathbb{E}\left[X\right]^2 \qquad \mathbb{E}\left[X\right] \text{ is just another constan}$$
$$= \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$$

Problem 6

$\text{Cov}\left[X, Y\right] = \mathbb{E}\left[X, Y\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$ But because
$X \perp Y \Rightarrow \mathbb{E}\left[X, Y\right] = \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right] \Rightarrow \text{Cov}\left[X, Y\right] = 0$

Problem 7

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2\right) dxdy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x^2 + y^2)\right) dxdy$$

$$= \int_{0}^{\infty} \int_{0}^{2\pi} \exp\left(-\frac{1}{2\sigma^2}\left[r^2 \cos^2(\theta) + r^2 \sin^2(\theta)\right]\right) rd\theta dr$$

$$= \int_{0}^{\infty} \int_{0}^{2\pi} \exp\left(-\frac{r^2}{2\sigma^2}\right) rd\theta dr$$

$$= 2\pi \int_{0}^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) rdr$$

$$= 2\pi \int_{0}^{\infty} \exp\left(-\frac{u}{2\sigma^2}\right) \frac{1}{2}du$$

$$= -2\pi\sigma^2 \exp\left(-\frac{u}{2\sigma^2}\right) |_{0}^{\infty}$$

$$= 2\pi\sigma^2$$

Now we just need to show that this normalizes the Gaussian. Take $y = x - \mu$ then

$$\mathcal{N}(x|\mu, \sigma^2) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}y^2\right) dy$$

Problem 8

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma)x\,dx$$

$$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y + \mu) \exp\left(-\frac{1}{2\sigma^2}y^2\right) dy \qquad y = x - \mu$$

Now split this into the sum of two integrals. The second integral has the form:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \mu \exp\left(-\frac{1}{2\sigma^2}y^2\right) dy$$

Which is just a normalized Gaussian times μ so this is just μ. Now since we are trying to prove that his equals μ I'm fairly confident that the left integral will go to zero. Let's try and prove this.

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \exp\left(-\frac{1}{2\sigma^2}y^2\right) dy$$

Well if we visualize this function it looks odd which would imply that the integral from $[-\infty, \infty]$ is 0. Let's prove that it is odd quickly. We want to show that $f(x) = -f(-x)$ where $f(x) = y \exp\left(-\frac{1}{2\sigma^2}y^2\right)$

$$-(-y) \exp\left(-\frac{1}{2\sigma^2}(-y)^2\right) = y \exp\left(-\frac{1}{2\sigma^2}y^2\right)$$

Thus or integral in question is odd and evaluates to 0 yielding μ as the desired result.

Next we need to show that:

$$\mathbb{E}\left[x^2\right] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma) x^2 dx = \mu^2 + \sigma^2$$

To do this differentiate both sides by σ^2 as follows:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx = 1$$

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx = \sqrt{2\pi\sigma^2}$$

$$\int_{-\infty}^{\infty} \frac{d}{d\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx = \frac{d}{d\sigma^2}\sqrt{2\pi\sigma^2}$$

$$\frac{1}{2\sigma^4} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) (x-\mu)^2 dx = \frac{\pi}{\sqrt{(2\pi\sigma^2)}}$$

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) (x-\mu)^2 dx = \frac{2\pi\sigma^4}{\sqrt{(2\pi\sigma^2)}}$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) (x-\mu)^2 dx = \sigma^2$$

$$\text{Var}\left[x\right] = \sigma^2$$

The left hand side of this equation is just the definition of variance. Now to complete the proof of (1.50) just use the alternate formulation of variance:

$$\mathbb{E}\left[(x-\mu)^2\right] = \mathbb{E}\left[x^2\right] - \mathbb{E}\left[x\right]^2$$
$$= \mathbb{E}\left[x^2\right] - \mu^2 \qquad \text{By part 1}$$
$$\sigma^2 = \mathbb{E}\left[x^2\right] - \mu^2 \qquad \text{By above}$$
$$\mathbb{E}\left[x^2\right] = \sigma^2 + \mu^2$$

Problem 9

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$
$$\frac{d\mathcal{N}(x|\mu,\sigma^2)}{dx} = -\frac{1}{\sigma^2}\mathcal{N}(x|\mu,\sigma^2)(x-\mu)$$
$$0 = \mathcal{N}(x|\mu,\sigma^2)(x-\mu)$$

But since $\mathcal{N}(x|\mu,\sigma^2)(x-\mu) > 0$ the only term that matters is $(x-\mu)$, which goes to 0 at $x=\mu$

The proof for the multivariate case is almost identical and is omitted

Problem 10

$$\log p(\mathbf{x}|\mu,\sigma^2) = -\frac{1}{\sqrt{2\pi\sigma^2}}\sum_{n=1}^{N}(x_n-\mu)^2 - \frac{N}{2}\log\sigma^2 - \frac{N}{2}\log(2\pi)$$

$$\frac{d\log p(\mathbf{x}|\mu,\sigma^2)}{d\mu} = -\frac{1}{\sqrt{\pi\sigma^2}}\sum_{n=1}^{N}(x_n-\mu)$$

$$0 = \sum_{n=1}^{N}(x_n-\mu)$$

$$= \sum_{n=1}^{N}x_n - \sum_{n=1}^{N}\mu$$

$$\mu = \sum_{n=1}^{N}x_n$$

Problem 11

Just take the derivative with respect to μ of our function:

$$\ln(x|\mu, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 - \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln(2\pi)$$

$$\frac{d}{d\mu} \ln(x|\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{n=1}^{N} x_n - \mu$$

$$0 = \sum_{n=1}^{N} x_n - \mu$$

$$\sum_{n=1}^{N} \mu = \sum_{n=1}^{N} x_n$$

$$N\mu = \sum_{n=1}^{N} x_n$$

$$\mu = \frac{1}{N} \sum_{n=1}^{N} x_n$$

Now repeat the same for σ^2

$$\frac{d}{d\mu}\ln(x|\mu,\sigma^2) = \frac{1}{2\sigma^4}\sum_{n=1}^{N}(x_n - \mu)^2 - \frac{N}{2\sigma^2}$$

$$0 = \frac{1}{2\sigma^4}\sum_{n=1}^{N}(x_n - \mu)^2 - \frac{N}{2\sigma^2}$$

$$\frac{N}{2\sigma^2} = \frac{1}{2\sigma^4}\sum_{n=1}^{N}(x_n - \mu)^2$$

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu)^2$$

Problem 17

The gamma function is defined as:

$$\Gamma(x) = \int_0^{\infty} u^{x-1}e^{-u}du$$

We use integration by parts with the following substitutions

$$u = u^{x-1} \qquad\qquad v = -e^{-u}$$
$$du = (x-1)u^{x-2}du \qquad\qquad dv = e^{-u}du$$

By integration by parts we have:

$$\Gamma(x) = -u^{x-1}e^{-u}|_0^\infty + (x-1)\int_0^\infty u^{x-2}e^{-u}du$$
$$= (x-1)\int_0^\infty u^{x-2}e^{-u}du$$
$$= (x-1)\Gamma(x-1)$$
$$= z\Gamma(z) \qquad\qquad\qquad z = x+1$$

Now see that $\Gamma(1) = 1$

$$\Gamma(1) = \int_0^\infty u^0 e^{-u}du$$
$$= \int_0^\infty e^{-u}du$$
$$= -e^{-u}|_0^\infty$$
$$= 1$$

To show that for all $x \in \mathbb{Z}^+$ $\Gamma(x+1) = x!$ we use an inductive argument. Our base case is $\Gamma(1) = 1$ which we have already done. Now let's assume that our inductive hypothesis is true and $\Gamma(x+1) = x!$. Then $\Gamma(x+2) = (x+1)\Gamma(x+1) = (x+1)x!$ and we are done.

Problem 18

We want to solve for S_D the surface area of a D dimensional sphere.

$$\prod_{i=1}^{D} \int_{-\infty}^\infty e^{-x_i^2} = S_D \int_0^\infty e^{-r^2} r^{D-1} dr$$

Let's solve the left hand side first. From problem 7 above we know that

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}x^2} = \sqrt{2\pi\sigma^2}$$

By setting $\sigma^2 = \frac{1}{2}$ we get that

$$\prod_{i=1}^{D} \int_{-\infty}^{\infty} e^{-x^2} = \prod_{i=1}^{D} \sqrt{\pi}$$
$$= \pi^{\frac{D}{2}}$$

Now let's look at the right hand side of the equation. We can tease this function into the Gamma function by making the substitution $u = r^2$ which implies that $dr = \frac{1}{2}u^{-1/2}$. With this insight we see that:

$$\int_0^{\infty} e^{-r^2} r^{D-1} dr = \frac{1}{2} \int_0^{\infty} e^{-u} u^{\frac{D}{2}-1} du$$
$$= \frac{1}{2}\Gamma(D/2)$$

Now just use simple algebra:

$$\pi^{\frac{D}{2}} = S_D \frac{1}{2}\Gamma(D/2)$$
$$S_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}$$

.

Next integrate the surface area with respect to the radius in order to get the equation for volume:

$$V_D = S_D \int_0^1 r^{D-1} dr$$

$$= S_D \left. \frac{r}{D} \right|_0^1$$

$$= \frac{S_D}{D}$$

Now to derive some common known volumes. For D =2:

$$\frac{2\pi^{D/2}}{D\Gamma(D/2)} = \frac{2\pi}{2\Gamma(1)}$$

$$= \pi$$

For D=3:

$$\frac{2\pi^{D/2}}{D\Gamma(D/2)} = \frac{2\pi^{3/2}}{3\Gamma(3/2)}$$

$$= \frac{2\pi^{3/2}}{3\frac{\sqrt{\pi}}{2}}$$

$$= \frac{4\pi}{3}$$

Problem 19