Advanced Image Analysis
Exercise

Rasmus Larsen

Revised January 28 2003 by H H Thodberg And S Darkner

Introduction

This exercise comprise four small subtasks, which you should be able to solve given the prerequisite courses for Advanced Image Analysis. For some tasks you may want to use Matlab. On the course homepage you will find a test image igal.mat for the exercise.

1 Filter design

In statistics we like to design optimal estimators of important aspects of the data. Well-known examples are the estimators of mean and standard deviation of a stochastic variable. This example shows how to apply this paradigm to image data. The simplest aspects we can think of for an image is the local mean and gradient, and we now want to estimate these three properties using a 3-by-3 neighbourhood. We want the estimator to be a linear function of the 9 pixel values because then we can express the estimation as a filter operation, which is implemented very efficiently in programs like Matlab or in libraries like the Intel Performance Library.

In this course we will regard the image as stochastic variables defined on a grid of pixels Thus consider an image function $f(x, y)$ sampled on a square grid and let the coordinates be $x \in \{-1, 0, 1\}$ and $y \in \{-1, 0, 1\}$. The sampled values in our neighourhood are denoted

$f(-1,1)$	$f(0,1)$	$f(1,1)$
$f(-1,0)$	$f(0,0)$	$f(1,0)$
$f(-1,-1)$	$f(0,-1)$	$f(1,-1)$

So forget about the image outside this neighbourhood. First we define what we want to estimate. To estimate the mean and gradients of f we approximate it by a plane within this 3×3 neighbourhood, and we parametrise the plane by

$$g(x, y) = a + b(x - x_0) + c(y - y_0)$$

Because our coordinate system is centered on the neighbourhood, we have $x_0 = y_0 = 0$, i.e.

$$g(x, y) = a + bx + cy$$

It is evident that b is a good approximation of $\frac{\partial f}{\partial x}$ and c is an approximation of $\frac{\partial f}{\partial y}$, and a approximates the mean, so we define the least square fit of f by g by the three variables a, b and c as the desired estimators. Thus we have 9 observations to derive three numbers which is a healthy problem. In your statistics course you have already seen a similar problem, namely multivariate linear regression (generalised linear model). Remeber that the three parameters are intertwined so they have to estimated jointly, and therefore we will organize our nine observations as a column vector, i.e. we get this model for our approximation

$$f = g + \epsilon = X\beta + \epsilon$$

where f is the image function values arranged as a column vector, g is the model values arranged as a column vector, and $\beta = [a, b, c]^T$. X are the coefficients to the parameters, and ϵ is an error term.

The least squares solution for the parameters, β, is found by minimizing

$$\|f - X\beta\|^2$$

with respect to β. Setting the diffential equal to zero yields

$$X^T f = \left[X^T X\right] \beta$$

1. Explain the above line of resoning to your mate, and write down the ~~8~~-by-3 matrix X.

2. Solve the equation for β and express it in terms of f. Arrange it as a three 3-by-3 filters - have you seen these filters before?

3. What are the optimal filter kernels if the squared errors of our 9 observations are weighted by the weight pattern shown below, which emphasises near pixels more than more remote? What are the conventional names of these filters? (Hint: If you remember weighted linear regression you may derive an equation involving a general diagonal weight matrix. Another way is to duplicate four of the observations and use the center observation 4 times. By the way, many students get this wrong, so before you give the answer check that your estimators deliver the expected results on 3-by-3 images generated by perfect planes with known slopes)

4. Explore filtering using the matlab convolution routine `conv2` or `filter2`

1	2	1
2	4	2
1	2	1

2 Fourier Transformation

Often it is advantageous to specify a filter in the Fourier domain. In order to construct the filter as a convolution kernel we would then employ the inverse Fourier transform.

We define the discrete Fourier transform $\mathcal{F}$ of f, where f is sampled at n points with the sampling distance 1 along the coordinate axis, as

$$\mathcal{F}(u) = \frac{1}{n}\sum_x f(x)\exp(-iux)$$

Evidently, $\mathcal{F}$ is periodic with the period 2π. Hence, $\mathcal{F}$ is completely defined by its values on an interval $\omega \leq u < \omega + 2\pi$. It is of no theoretical consideration where such an interval is positioned.

However, for filter design it is practical to position the spatial as well as the Fourier domain around origo. For functions with an even number of samples we choose to position origo in the left of the two middle samples.

1. For n samples, where n is an odd number, what are the sampling points in the spatial and the Fourier domain?

Let the (column) vector of sample points in the spatial domain be denoted x and the vector of sample points in the Fourier domain be denoted u.

Furthermore, let the column vector of the sampled values of the function f be denoted x_f and let the column vector of the sampled values of $\mathcal{F}$ be denoted x_F. The relation between x_f and x_F is given by

$$x_F = Bx_f$$

1. What are the values of B?
2. Show that B is an orthogonal (but not orthonormal) basis.
3. Implement a matlab function that computes B as a function of odd n.

Hence, we have shown that the Fourier transform is nothing but a change of base.

3 Linear Transformations

The Landsat TM 4 satellite has an imaging spectrometer that measures the sun reflection in 4 frequency bands (G, R, IR1, IR2). Copy the image igal.mat (Landsat/MSS images of the Igaliko fjord in South Greenland) availabe on the course home page to your current matlab directory and inspect the image using the commands

```
load igal;
imagesc([igal(:,:,1) igal(:,:,2); igal(:,:,3) igal(:,:,4)]);
colormap gray;
axis equal;
axis off;
```

Let these four variables be denoted $X = (X_1, \ldots, X_4)^T$.

Let the pixelwise (estimated) mean reflectance and (estimated) dispersion of the reflectance be μ and Σ (with diagonal elements σ_i^2 and off-diagonal elements σ_{ij}, where i is row number and j is column number), respectively.

We want to find the linear combination of the original spectral bands $(X - \mu)$ that contains most of the variation.

For another image of the same region we got

$$\mu = \begin{bmatrix} 15.476028 \\ 13.710770 \\ 19.574310 \\ 10.387772 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 4.618911 & 4.711057 & 0.915911 & -0.718644 \\ 4.711057 & 21.436799 & 39.920608 & 28.907793 \\ 0.915911 & 39.920608 & 107.314780 & 84.619395 \\ -0.718644 & 28.907793 & 84.619395 & 70.643604 \end{bmatrix}$$

The eigenvalues of Σ are (represented as a diagonal matrix)

$$\Lambda = \begin{bmatrix} 189.956573 & & & \\ & 11.236966 & & \\ & & 2.009580 & \\ & & & 0.810975 \end{bmatrix}$$

The eigenvectors of Σ are (shown in the columns)

$$P = \begin{bmatrix} 0.008513 & 0.566824 & 0.608954 & 0.554809 \\ 0.280628 & 0.739548 & -0.152479 & -0.592509 \\ 0.749509 & 0.006975 & -0.455007 & 0.480787 \\ 0.599509 & -0.362948 & 0.631580 & -0.331608 \end{bmatrix}$$

1. Recompute the same quantities but now for the provided image
2. What is the linear combination Y of $X - \mu$ that explains most of the variation? What is this linear combination called?
3. Write an analytic expression for the correlations between this linear combination and the 4 original spectral bands, and then compute the actual correlations both directly and using your derived expression

4 Classification

Let us assume that pixels belong to one of two classes, π_1 or π_2. Let the prior probabilities for the two classes be

$$P\{C = \pi_1\} = \frac{1}{3}$$
$$P\{C = \pi_2\} = \frac{2}{3}$$

Let the image be corrupted by noise, and let the noise be class dependent, i.e.

$$X \mid C = \pi_1 \in N\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\right)$$
$$X \mid C = \pi_2 \in N\left(\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

The density of the two dimensional normal distribution is

$$f(x) = \frac{1}{2\pi}\frac{1}{\sqrt{\det \Sigma}} \exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$

1. Assuming equal losses from misclassification, what is the curve that separates the two populations in feature space?