Sample Math Used in CS4243

The mathematics used in CS4243 are basically linear algebra, calculus, and probability. This summary contains a sampling of the math used.

1. Linear Algebra

- An element \( \mathbf{v} \) in an \( n \)-dimensional vector space \( V \) can be represented as a linear combination of \( n \) basis vectors:

\[
\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{e}_i
\]

(1)

where \( \mathbf{e}_i \) is a basis vector and \( a_i \) is the corresponding coefficient.

- Given a set of \( n \) vector \( \mathbf{x}_i \), the mean of the vectors is

\[
\mathbf{m} = E\{\mathbf{x}\} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i
\]

(2)

and the covariance matrix is

\[
\mathbf{C} = E\{(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^T\} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T - \mathbf{m} \mathbf{m}^T
\]

(3)

- An eigensystem relates a matrix \( \mathbf{A} \) and a set of vectors \( \mathbf{e}_i \) by the equations

\[
\mathbf{A} \mathbf{e}_i = \lambda_i \mathbf{e}_i
\]

(4)

The vectors \( \mathbf{e}_i \) are the eigenvectors and the corresponding \( \lambda_i \) are the eigenvalues.

- Consider a set of simultaneous equations

\[
\mathbf{A} \mathbf{x} = \mathbf{b}
\]

(5)

where \( \mathbf{A} \) is an \( N \times N \) matrix, \( \mathbf{b} \) and \( \mathbf{x} \) are vectors. This equation defines \( \mathbf{A} \) as a mapping from the vector space \( \mathbf{x} \) to the vector space \( \mathbf{b} \).

  - If \( \mathbf{A} \) is singular, then there is some subspace \( \mathbf{x} \) that is mapped to zero, i.e., \( \mathbf{A} \mathbf{x} = \mathbf{0} \). This subspace is called the nullspace, and its dimensionality is called the nullity of \( \mathbf{A} \).
  
  - The non-null subspace that is mapped from \( \mathbf{x} \) is called the range of \( \mathbf{A} \) and the dimensionality of the range is called the rank of \( \mathbf{A} \).
  
  - So, the rank plus the nullity of \( \mathbf{A} \) equals \( N \).
A 3D point $W$ with world coordinate $X = (X, Y, Z)^T$ is mapped onto the camera coordinate system at $X^c = (X^c, Y^c, Z^c)^T$ by a rigid transformation

$$X^c = RX + T$$

where $R = (R_1^T, R_2^T, R_3^T)^T$ is the $3 \times 3$ rotation matrix and $T = (T_X, T_Y, T_Z)^T$ is the translation vector that relate the camera and the world coordinate frames.

The rotation matrix can be specified in terms of three Euler angles, pitch (vertical angle) $\omega$, yaw (horizontal angle) $\phi$, and roll $\kappa$:

$$R = \begin{bmatrix}
\cos \phi \cos \kappa & \sin \omega \sin \phi \cos \kappa + \cos \omega \sin \kappa & -\cos \omega \sin \phi \cos \kappa + \sin \omega \sin \kappa \\
-\cos \phi \sin \kappa & -\sin \omega \sin \phi \sin \kappa + \cos \omega \cos \kappa & \cos \omega \sin \phi \sin \kappa + \sin \omega \cos \kappa \\
\sin \phi & -\sin \omega \cos \phi & \cos \omega \cos \phi
\end{bmatrix}$$

which is an orthonormal matrix, i.e.,

$$R^T R = I$$

where $I$ is the identity matrix.

2. Calculus

- The gradient $\nabla f$ of a function $f$ of two variables $x$ and $y$ is the first partial derivative with respect to $x$ and $y$

$$\nabla f(x, y) = \begin{bmatrix}
\frac{\partial f}{\partial x} \\
\frac{\partial f}{\partial y}
\end{bmatrix}.$$  

The magnitude of the gradient is

$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}.$$  

- The Laplacian $\nabla^2 f$ of a function $f$ of two variables $x$ and $y$ is

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$  

- The image of a moving object changes over time. Assuming small motion, the intensity $E(x, y, t)$ at location $(x, y)$ will remain constant, so that

$$\frac{\partial E}{\partial t} = 0.$$  

Using the chain rule for differentiation, we see that

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0.$$
3. Probability

- The notation \( P(H|E) \) denote the probability that \( H \) occurs given that evidence \( E \) has occurred. This conditional probability is defined by the joint probability \( P(H, E) \), that is, the probability that both \( H \) and \( E \) occur:

\[
P(H|E) = \frac{P(H, E)}{P(E)} .
\]  
(14)

- From the definition of conditional probability, we have

\[
P(E|H) = \frac{P(E, H)}{P(H)} = \frac{P(H, E)}{P(H)} .
\]  
(15)

Therefore,

\[
P(H|E) P(E) = P(E|H) P(H)
\]  
(16)

or

\[
P(H|E) = \frac{P(E|H) P(H)}{P(E)}
\]

which is known as the Bayes rule.

- The entropy \( E \) of a probability distribution \( P(i) \) is

\[
E = - \sum_i P(i) \log P(i) .
\]  
(18)