COMP4121 Lecture Notes - Topic 10

The Discrete Fourier Transform, the Discrete Cosine Transform and JPEG

LiC: Aleks Ignjatovic

ignjat@cse.unsw.edu.au

THE UNIVERSITY OF
NEW SOUTH WALES

School of Computer Science and Engineering
The University of New South Wales
Sydney 2052, Australia
1 DFT as a change of basis

Recall that the scalar product (also called the dot product) of two vectors with real coordinates, \( \mathbf{x} = (x_0, x_1, \ldots, x_{n-1}) \) and \( \mathbf{y} = (y_0, y_1, \ldots, y_{n-1}) \), \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^{n-1} \), denoted by \( \langle \mathbf{x}, \mathbf{y} \rangle \) (or \( \mathbf{x} \cdot \mathbf{y} \)) is defined as

\[
\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^{n-1} x_i y_i.
\]

If the coordinates of our vectors are complex numbers, i.e., if \( \mathbf{x}, \mathbf{y} \in \mathbb{C}^n \), then the scalar product of such two vectors is defined as

\[
\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^{n-1} x_i \bar{y}_i,
\]

where \( \bar{z} \) denotes the complex conjugate of \( z \), i.e., \( a + ib = a - ib \).

Such a scalar product also defines the norm of a vector by

\[
\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.
\]

Geometrically, the norm plays the role of the length of a vector (and, in case of \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \), it is just the length of the vector).

The usual basis of both \( \mathbb{R}^n \) and \( \mathbb{C}^n \) (depending on whether the scalars of the vector space are real or complex numbers) is given by

\[
\mathcal{B} = \{(1,0,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,0,\ldots,1)\}
\]

We obviously have for any vector \( \mathbf{a} = (a_0, a_1, a_2, \ldots, a_{n-1}) \),

\[
(a_0, a_1, a_2, \ldots, a_{n-1}) = a_0(1,0,0,\ldots,0) + a_1(0,1,0,\ldots,0) + \ldots + a_{n-1}(0,0,\ldots,1)
\]

Let us set

\[
\mathbf{e}_0 = (1,0,0,\ldots,0); \quad \mathbf{e}_1 = (0,1,0,\ldots,0); \quad \ldots \quad \mathbf{e}_{n-1} = (0,0,\ldots,1);
\]

Thus,

\[
(a_0, a_1, a_2, \ldots, a_{n-1}) = a_0\mathbf{e}_0 + a_1\mathbf{e}_1 + \ldots + a_{n-1}\mathbf{e}_{n-1} = \sum_{m=0}^{n-1} a_m \mathbf{e}_m
\]

Let us denote the complex number \( e^{i \frac{2\pi}{n}} \) by \( \omega_n \); such a number is a primitive root of unity because \( (\omega_n)^n = \omega_n^m = 1 \), and the set of all complex numbers \( z \) which satisfy \( z^n = 1 \) is precisely the set of \( n \) powers of \( \omega_n \), i.e., \( z^n = 1 \) if and only if \( z \) is of the form \( z = (\omega_n)^k \) for some \( k \) such that \( 0 \leq k \leq n - 1 \).

We now introduce another basis, this time only in \( \mathbb{C}^n \), given by \( \mathcal{F} = \{\mathbf{f}_0, \ldots, \mathbf{f}_{n-1}\} \), where

\[
\mathbf{f}_k = (\omega_n^0, \omega_n^1, \ldots, \omega_n^{n-1}) = (1, \omega_n, \ldots, \omega_n^{n-1});
\]

Thus, the coordinates of \( \mathbf{f}_k \) are the first \( n \) powers of the \( k \)th power of \( \omega_n \).

To show that this is indeed a basis we have to show that these vectors are linearly independent. In fact, they form an orthogonal basis. Two vectors are mutually orthogonal if \( \langle \mathbf{x}, \mathbf{y} \rangle = 0 \). Prove as a homework that if a set of \( n \) vectors in \( \mathbb{C} \) are pairwise mutually orthogonal, then they must be linearly independent.

To see that vectors \( \mathbf{f}_k \) and \( \mathbf{f}_m \) are orthogonal if \( k \neq m \) we compute their scalar product:

\[
\langle \mathbf{f}_k, \mathbf{f}_m \rangle = \sum_{p=0}^{n-1} (\omega_n^k)^p (\omega_n^m)^{-p} = \sum_{p=0}^{n-1} (\omega_n^k)^p (\omega_n^{-m})^p = \sum_{p=0}^{n-1} (\omega_n^{-m})^{-p} = \sum_{p=0}^{n-1} (\omega_n^{-k+m})^p
\]

The last sum is a sum of a geometric progression with ratio \( q = \omega_n^{k-m} \) and thus, using formula

\[
\sum_{p=0}^{n-1} q^p = \frac{1-q^n}{1-q}
\]

we get

\[
\langle \mathbf{f}_k, \mathbf{f}_m \rangle = \frac{1-(\omega_n^{k-m})^n}{1-\omega_n^{k-m}} = 0
\]
because \((\omega_n^{k-m})^n = (\omega_n^m)^{k-m} = 1\) (the denominator is different from 0 because we have assumed that \(k \neq m\)). Let us compute the norm of these vectors:

\[
\|\tilde{f}_k\|^2 = (\tilde{f}_k, \tilde{f}_k) = \sum_{p=0}^{n-1} (\omega_n^k)^p (\omega_n^m)^{-k} = \sum_{p=0}^{n-1} (\omega_n^k)^p (\omega_n^m)^{-k} = \sum_{p=0}^{n-1} (\omega_n^k)^0 = \sum_{p=0}^{n-1} 1 = n.
\]

Thus, \(\|\tilde{f}_k\| = \sqrt{n}\); consequently, if we define vectors \(\phi_k = \tilde{f}_k / \sqrt{n}\), then these \(n\) vectors form an orthonormal basis, \(\Phi = \{\phi_0, \ldots, \phi_{n-1}\}\).

Recall that in \(\mathbb{R}^2\) we have \((x, y) = \|x\| \cdot \|y\| \cdot \cos(\angle x, y)\). Thus, if \(y\) is a unit vector, \(\|y\| = 1\), then we have \((x, y) = \|x\| \cdot \cos(\angle x, y)\), i.e., \((x, y)\) is just the length of the orthogonal projection of \(x\) onto the direction of \(y\):

\[
\langle x, y \rangle = \langle x, e_1 \rangle e_1 + \langle x, e_2 \rangle e_2 + \langle x, e_3 \rangle e_3.
\]

Let us now calculate the coefficient vector \(\tilde{c} = (\langle c, \phi_m \rangle)_m\); using definitions of \(\phi_m\) and of scalar product in \(\mathbb{C}^n\) we have

\[
\tilde{c} = (\langle c, \phi_m \rangle)_m = \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} c_k e^{-i \frac{2\pi}{n} m k} \right)_m
\]

The sequence (i.e., vector) \(\tilde{c}\) is called the Discrete Fourier Transform (DFT) of sequence \(c\). Thus, the DFT of a vector \(c = (c_0, \ldots, c_{n-1})\) is nothing but the sequence \(\tilde{c} = (\tilde{c}_0, \ldots, \tilde{c}_{n-1})\) of the coordinates of \(c\) in the orthonormal basis \(\Phi = \{\phi_0, \ldots, \phi_{n-1}\}\).

But why would we consider such a complicated basis? Let us consider complex sinusoids of the form:

\[
\text{Si}_k(t) = \frac{1}{\sqrt{n}} \left( \cos \left( \frac{2\pi}{n} k \cdot t \right) + i \sin \left( \frac{2\pi}{n} k \cdot t \right) \right) = e^{i \frac{2\pi}{n} k t}.
\]
of frequencies $2\pi k/n$, for all $0 \leq k \leq n - 1$. Then each of the vectors $\varphi_k$ is just a sequence of samples of these functions, evaluated at integers $0, 1, \ldots, n - 1$:

$$\varphi_k = (\text{Si}_k(p) : p = 0, 1, \ldots, n - 1)$$

Thus, if we represent a vector $x$ in such a basis, i.e., as $x = \sum_{k=0}^{n-1} \hat{x}_k \varphi_k$, we have represented $x$ as a linear combination of samples of such complex sinusoids. Namely if we see the sequence $x$ as a sequence of samples of a function $x(t)$, then

$$x(t) \approx \sum_{k=0}^{n-1} \hat{x}_k \text{Si}_k(t) = \sum_{k=0}^{n-1} \hat{x}_k e^{i \frac{2\pi k}{n} \cdot t}$$

where the equality is exact on integers $0, 1, \ldots, n - 1$. Let us write each coordinate $\hat{x}_k$ in the polar form:

$$\hat{x}_k = |\hat{x}_k| e^{i \arg(\hat{x}_k)}$$

i.e., we have done, what is called a spectral analysis of $x$, because the values $|\hat{x}_k|/\sqrt{n}$ represent the amplitudes of the complex sinusoids and the arguments $\arg(\hat{x}_k)$ represent the phase shifts of these complex sinusoids. Note that the equality is exact only on integers $0, \ldots, n - 1$, i.e., for $m$ such that $0 \leq m \leq n - 1$, we have

$$x(m) = \sum_{k=0}^{n-1} |\hat{x}_k| e^{i \frac{2\pi k}{n} \cdot m \cdot t + \arg(\hat{x}_k)}$$

But what happens if the samples $x_0, x_1, \ldots x_{n-1}$ come from a signal containing frequencies different from those of the form $2\pi k/n$? To understand this please look at the Mathematica file available at [http://www.cse.unsw.edu.au/~cs4121/DFT_and_DCT.nb](http://www.cse.unsw.edu.au/~cs4121/DFT_and_DCT.nb)