LECTURE #6: Mean-Value Theorems
Appendix D (Mean-Value Theorems, pp. 744-748):

Setting

Let \( g(z) \in \mathbb{R} \) denote a \textit{real-valued} function of a possibly vector argument \( z \). We assume that \( g(z) \) is differentiable whenever necessary. In this appendix, we review useful integral equalities that involve increments of the function \( g(z) \) and increments of its gradient vector; the equalities correspond to extensions of the classical mean-value theorem from single-variable real calculus to the case of functions of several and possibly complex variables. We shall use the results of this appendix to establish useful bounds on the increments of strongly convex functions later in Appendix E. We again treat both cases of real and complex arguments.
Real Arguments
Real Arguments

Consider first the case in which the argument $z \in \mathbb{R}^M$ is real-valued. We pick any $M$–dimensional vectors $z_0$ and $\Delta z$ and introduce the following real-valued and differentiable function of the scalar variable $t \in [0, 1]$

$$f(t) \triangleq g(z_0 + t \Delta z) \quad (D.1)$$
Then, it holds that

\[ f(0) = g(z_0), \quad f(1) = g(z_0 + \Delta z) \]  \hspace{1cm} (D.2)

Using the fundamental theorem of calculus (e.g., [36, 151]) we have:

\[ f(1) - f(0) = \int_0^1 \frac{df(t)}{dt} dt \]  \hspace{1cm} (D.3)

It further follows from definition (D.1) that

\[ \frac{df(t)}{dt} = \frac{d}{dt} \left[ g(z_0 + t \Delta z) \right] = \left[ \nabla_z g(z_0 + t \Delta z) \right] \Delta z \]  \hspace{1cm} (D.4)
in terms of the inner product computation on the far right, where $\nabla_z g(z)$ denotes the (row) gradient vector of $g(z)$ with respect to $z$. Substituting (D.4) into (D.3) we arrive at the first desired mean-value theorem result (see, e.g., [191]):

$$g(z_o + \Delta z) - g(z_o) = \left( \int_0^1 \nabla_z g(z_o + t\Delta z) dt \right) \Delta z$$  \hspace{1cm} (D.5)
Real Arguments

This result is a useful equality and it holds for any differentiable \textit{(not necessarily convex)} real-valued function \( g(z) \). The expression on the right-hand side is an inner product between the column vector \( \Delta z \) and the result of the integration, which is a row vector. Expression (D.5) tells us how the increment of the function \( g(z) \) in moving from \( z = z_o \) to \( z = z_o + \Delta z \) is related to the integral of the gradient vector of \( g(z) \) over the segment \( z_o + t \Delta z \) as \( t \) varies over the interval \( t \in [0, 1] \).
Real Arguments

We can derive a similar relation for increments of the gradient vector itself. To do so, we introduce the column vector function $h(z) = \nabla_{z^T} g(z)$ and apply (D.5) to its individual entries to conclude that

$$h(z_o + \Delta z) - h(z_o) = \left( \int_0^1 \nabla_z h(z_o + r \Delta z) dr \right) \Delta z \tag{D.6}$$

Replacing $h(z)$ by its definition, and transposing both sides of the above equality, we arrive at another useful mean-value theorem result:

$$\nabla_z g(z_o + \Delta z) - \nabla_z g(z_o) = \Delta z^T \left( \int_0^1 \nabla^2 \nabla_z g(z_o + r \Delta z) dr \right) \tag{D.7}$$
Real Arguments

This expression tells us how increments in the gradient vector in moving from $z = z_0$ to $z = z_0 + \Delta z$ are related to the integral of the Hessian matrix of $g(z)$ over the segment $z_0 + r \Delta z$ and $r$ varies over the interval $r \in [0, 1]$. In summary, we arrive at the following statement.
**Lemma D.1** (Mean-value theorem: Real arguments). Consider a real-valued and twice-differentiable function \( g(z) \in \mathbb{R} \), where \( z \in \mathbb{R}^M \) is real-valued. Then, for any \( M \)-dimensional vectors \( z_o \) and \( \Delta z \), the following increment equalities hold:

\[
g(z_o + \Delta z) - g(z_o) = \left( \int_0^1 \nabla_z g(z_o + t \Delta z) dt \right) \Delta z \quad \text{(D.8)}
\]

\[
\nabla_z g(z_o + \Delta z) - \nabla_z g(z_o) = (\Delta z)^T \left( \int_0^1 \nabla^2_z g(z_o + r \Delta z) dr \right) \quad \text{(D.9)}
\]
Complex Arguments
Complex Arguments

We now extend results (D.8) and (D.9) to the case when $z \in \mathbb{C}^M$ is complex valued. The extension can be achieved by replacing $z = x + jy$ by its real and imaginary parts $\{x, y\}$, applying results (D.8) and (D.9) to the resulting function $g(v)$ of the $2M \times 1$ extended real variable

$$v = \text{col}\{x, y\}$$

(D.10)

and then transforming back to the complex domain. Indeed, as remarked earlier in (C.25), it is straightforward to verify that the vector $v$ is related to the vector

$$u \triangleq \text{col}\{z, (z^*)^T\}$$

(D.11)
Complex Arguments

as follows:

\[
\begin{align*}
\begin{bmatrix}
    z \\ (z^*)^T
\end{bmatrix} & \triangleq u \\
\begin{bmatrix}
    x \\ y
\end{bmatrix} & \triangleq v
\end{align*}
\]

\[
\begin{align*}
\begin{bmatrix}
    I_M & jI_M \\
    I_M & -jI_M
\end{bmatrix} \begin{bmatrix}
    x \\ y
\end{bmatrix} & \triangleq D
\end{align*}
\]

\[
\begin{align*}
\begin{bmatrix}
    1 \\
    2
\end{bmatrix} \begin{bmatrix}
    I_M & I_M \\
    -jI_M & jI_M
\end{bmatrix} \begin{bmatrix}
    z \\ (z^*)^T
\end{bmatrix} & = \frac{1}{2} D^*
\end{align*}
\]

(D.12)
Complex Arguments

or, more compactly,

\[ u = Dv \quad \text{and} \quad v = \frac{1}{2} D^* u \]  \hspace{1cm} (D.13)

where we used the fact from (B.28) that \( DD^* = 2I_{2M} \). We can now apply (D.8) to \( g(v) \) to get

\[ g(v_o + \Delta v) - g(v_o) = \left( \int_0^1 \nabla_v g(v_o + t\Delta v) dt \right) \Delta v \]  \hspace{1cm} (D.14)
Complex Arguments

where $\nabla_v g(v)$ denotes the gradient vector of $g(v)$. We can rewrite (D.14) in terms of the original complex variables $\{z_o, \Delta z\}$. To do so, we call upon relation (C.32) and the equality $g(z) = g(v)$ to rewrite (D.14) as

$$g(z_o + \Delta z) - g(z_o) = \quad \text{(D.15)}$$

\[
\begin{align*}
& \overset{\text{(D.13)}}{=} \frac{1}{2} \left( \int_0^1 \nabla_v g(v_o + t \Delta v) dt \right) D^* \begin{bmatrix} \Delta \Delta v \\
\end{bmatrix} \\
& \overset{\text{(C.32)}}{=} \left( \int_0^1 \begin{bmatrix} \nabla_z g(z_o + t \Delta z) \\
(\nabla_{z^*} g(z_o + t \Delta z))^T 
\end{bmatrix} dt \right) \begin{bmatrix} \Delta z \\
(\Delta z^*)^T 
\end{bmatrix}
\end{align*}
\]
Complex Arguments

We then arrive at the desired mean-value theorem result in the complex case:

\[ g(z_o + \Delta z) - g(z_o) = 2 \text{Re} \left\{ \left( \int_0^1 \nabla_z g(z_o + t\Delta z) \, dt \right) \Delta z \right\} \quad (D.16) \]

where we used the fact that for real-valued functions \( g(z) \) it holds that

\[ \nabla_z^* g(z) = [\nabla_z g(z)]^* \quad (D.17) \]

Expression (D.16) is the extension of (D.8) to the complex case.
Similarly, applying (D.6) to \( h(v) = \nabla_{v^T} g(v) \) we obtain that for any \( v_o \) and \( \Delta v \):

\[
\nabla_{v^T} g(v_o + \Delta v) - \nabla_{v^T} g(v_o) = \left( \int_0^1 \nabla^2_{v} g(v_o + r \Delta v) dr \right) \Delta v \quad (D.18)
\]

Multiplying from the left by \( \frac{1}{2} D \) and using (C.30)–(C.31), as well as the fact that \( \frac{1}{4}DH_v(v)D^* = H_c(u) \) (recall (B.26)), we find that relation (D.18) defined in terms of \( \{v_o, \Delta v\} \) can be transformed into the mean-value theorem relation (D.20) in terms of the variables \( \{z_o, \Delta z\} \).
Complex Arguments

Expression (D.20) is the extension of (D.9) to the complex case. Observe how both gradient vectors relative to $z^*$ and $z^T$ now appear in the relation. We show below in Example D.1 how the relation can be simplified in the special case when the Hessian matrix turns out to be block diagonal. In summary, we arrive at the following result.
**Lemma D.2** (Mean-value theorem: Complex arguments). Consider a real-valued and twice-differentiable function $g(z) \in \mathbb{R}$, where $z \in \mathbb{C}^M$ is complex-valued. Then, for any $M$-dimensional vectors $z_o$ and $\Delta z$, the following increment equalities hold:

\[
g(z_o + \Delta z) - g(z_o) = 2 \text{Re} \left\{ \left( \int_0^1 \nabla_z g(z_o + t \Delta z) dt \right) \Delta z \right\} \quad \text{(D.19)}
\]

\[
\begin{bmatrix}
\nabla_{z^*} g(z_o + \Delta z) \\
\nabla_{\bar{z}} g(z_o + \Delta z)
\end{bmatrix} - \begin{bmatrix}
\nabla_{z^*} g(z_o) \\
\nabla_{\bar{z}} g(z_o)
\end{bmatrix} = \left( \int_0^1 \nabla^2_z g(z_o + r \Delta z) dr \right) \begin{bmatrix}
\Delta z \\
(\Delta z^*)^T
\end{bmatrix} \quad \text{(D.20)}
\]
Example D.1 (Block diagonal Hessian matrix). Consider the real-valued quadratic function

$$g(z) = \kappa + a^* z + z^* a + z^* C z$$  \hspace{1cm} (D.21)

where $\kappa$ is a real scalar, $a$ is a column vector, and $C$ is a Hermitian matrix. Then, the Hessian matrix of $g(z)$ is block diagonal and given by

$$\nabla^2_z g(z) \equiv H_c(u) = \begin{bmatrix} C & 0 \\ 0 & C^T \end{bmatrix}$$ \hspace{1cm} (D.22)
Example #D.1

In this case, expression (D.20) decouples into two separate and equivalent relations. Keeping one of the relations we get

$$\nabla_z g(z_o + \Delta z) = \nabla_z g(z_o) + (\Delta z)^* C$$  \hspace{1cm} (D.23)

Obviously, in this case, this relation could have been deduced directly from expression (D.21) by using the fact that

$$\nabla_z g(z) = a^* + z^* C$$  \hspace{1cm} (D.24)
End of Lecture