Notes and Comments on Lecture 3

We use similar notation here as in other optimization courses (e.g., Nonlinear optimization), see the following page.

On partial derivatives. Matrix notation when writing gradients, Jacobians and Hessians.

Pages 2-7

Suppose $x, \dot{x}$ have $n+m$ components, f denotes n constraint equations, p has n components. Then,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_{n+m} \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad p = \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}.$$

Page 5

$g_x, g_{\dot{x}}$ are gradients of g with respect to $x, \dot{x}$. E.g.,

$$g_x = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \vdots \\ \frac{\partial g}{\partial x_{n+m}} \end{bmatrix}, \quad (n+m) \times 1 \text{ matrix; } g_x^\top \text{ its transpose}$$

$$f_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_{n+m}} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_{n+m}} \end{bmatrix}, \quad n \times (n+m) \text{ Jacobian matrix of } f$$

Thus we have: $\underset{1\times(n+m)}{g_x^\top} + \underset{n\times(n+m)}{\overset{1\times n}{p^\top} f_x}$; the sum is a $1 \times (n+m)$ matrix. Similarly, δx is an $(n+m) \times 1$ column matrix, and so on.

Page 8

For optimal control problems, we assume:

- x is an n-dimensional column vector,
- u is an m-dimensional column vector,
- f is an n-dimensional column vector,

and so on.

Notation in Optimization Courses

Vectors in $\mathbb{R}^n, X, C$ and so on are written as column vectors in matrix notation:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \quad x^\top y = \sum_{i=1}^{n} x_i y_i \quad \text{is the dot product.}$$

If $f : \mathbb{R}^n \to \mathbb{R}$, we define the gradient of $f(x)$ as a column vector:

$$\nabla f(x) := f_x(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

If, e.g., $f(x) = c^\top x$, c a constant vector, then $\nabla f(x) = c$. If $f(x) = c^\top Ax$, $c \in \mathbb{R}^m$ constant, A a constant $m \times n$ matrix, then $\nabla f(x) = A^\top c$.

The Jacobian matrix of a function $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$, is defined as an $m \times n$ matrix,

$$J^f(x) = f_x(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

If, e.g., $f(x) = Ax$, A an $m \times n$ constant matrix, then $f_x(x) = A$ is a constant matrix.

Let $f : \mathbb{R}^n \to \mathbb{R}^n$. The Hessian matrix of f at x is defined as $H^f(x) = f_{xx}(x) = \left[\delta f(x)/\delta x_i \delta x_j\right]$, an $n \times n$ matrix, usually symmetric, $H^f(x)^\top = H^f(x)$.

Let $f(x) = x^\top Qx$, $Q^\top = Q$ is a constant $n \times n$ matrix. Then,

$$f_x(x) = Q^\top x + Qx = 2Qx; \; H^f(x) = f_{xx}(x) = 2Q \quad \square$$