Appendix A Elements of Mathematics: A Quick Reference

He calls to me out of Seir, Watchman, what of the night? Watchman, what of the night? The watchman said, The morning comes, and also the night; if you will inquire, inquire, and come again. (Isaiah 21:11–12)
The people to whom this was said has enquired and tarried for more than two millennia, and we are shaken when we realize its fate. ... Nothing is gained by yearning and tarrying alone, ... We shall set to work and meet the 'demands of the day,'... if each finds and obeys the demon who holds the fibers of his very life.

—Max Weber (1918), *Speech at Ludwig-Maximilians-Universität München*

*

1 Elementary Mathematics

The Jensen concavity inequality For a concave function $f(x) : \mathbb{R} \to \mathbb{R}$ and $m = 1, 2, \ldots$, then

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \geq \sum_{i=1}^{m} \lambda_i f(x_i) \text{ with } x_i \in \mathbb{R}, \lambda_i \geq 0, \text{ and } \sum_{i=1}^{m} \lambda_i = 1.$$

Example: *The cummulative distribution function $F(x)$ with stochastic consumption pay-off x is called a lottery $F(x)$. Then the neoclassical (strictly concave) von Neumann-Morgenstern utility from the lottery satisfies*

$$\int u(x) dF(x) < u\left(\int x dF(x)\right), \ i.e. \mathbb{E}\left[u(x)\right] < u\left(\mathbb{E}\left[x\right]\right).$$

Further, a lottery $F(\cdot)$ first-order stochastically dominates another lottery $G(\cdot)$ when

$$\int u(x) dF(x) \geq \int u(x) dG(x),$$

and such first-order stochastical domination holds if and only if $F(x) \leq G(x)$ for all x. If the two lotteries have the same mean, then $F(\cdot)$ second-order stochastically dominates $G(\cdot)$ when

$$\int u(x) dF(x) \geq \int u(x) dG(x).$$

2 Analysis

Euler's formula of homogenous functions Suppose $F(x_1, x_2, \ldots, x_n)$ is homogeneous of degree r and differentiable. Then at any $(x_1, x_2, \ldots, x_n)$

$$\frac{\partial F(x_1, x_2, \ldots, x_n)}{\partial x_1} x_1 + \frac{\partial F(x_1, x_2, \ldots, x_n)}{\partial x_2} x_2 + \ldots + \frac{\partial F(x_1, x_2, \ldots, x_n)}{\partial x_n} x_n = rF(x_1, x_2, \ldots, x_n).$$

Proof: *By definition for arbitrary $\lambda > 0$*

$$F(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) - \lambda^r F(x_1, x_2, \ldots, x_n) = 0.$$

Differentiate it with respect to λ

$$\sum_{i=1}^{n} \frac{\partial F(\lambda x_1, \lambda x_2, \ldots, \lambda x_n)}{\partial(\lambda x_i)} x_i = r\lambda^{r-1} F(x_1, x_2, \ldots, x_n).$$

Since λ is arbitrarily taken, the equation above surely holds when $\lambda = 1$. □

Geometric series The sum of the sequence $a, aq, aq^2, \ldots, aq^n$ is

$$\sum_{i=0}^{n} aq^i = \frac{a\left(1 - q^{n+1}\right)}{1 - q}. \text{ With } q \in (0, 1), \lim_{n\to+\infty} \sum_{i=0}^{n} aq^i = \frac{a}{1 - q}.$$

Proof: *Define*

$$\begin{aligned} S &= \sum_{i=0}^{n} aq^i = a + aq + aq^2 + \ldots + aq^n, \text{ and} \\ Sq &= q \sum_{i=0}^{n} aq^i = aq + aq^2 + \ldots + aq^n + aq^{n+1}. \text{ Then} \\ S - Sq &= a - aq^{n+1}, \\ S &= \frac{a\left(1 - q^{n+1}\right)}{1 - q}. \end{aligned}$$

□

The intermediate value theorem (Theorem of Balzano) If the function $f(x) : [a, b] \to \mathbb{R}$ is continuous and $[f(a) - c] [f(b) - c] \leq 0$ ($c \in \mathbb{R}$), then the equation $f(x) = c, x \in [a, b]$ has a solution. That is, if the two ends of the function, $f(a)$ and $f(b)$ lie on different sides of c, then the continuous curve $f(x)$ must cross c at least once somewhere between a and b.

L'Hôpital's rule Suppose that $f(x) : \mathbb{R} \to \mathbb{R}$ and $g(x) : \mathbb{R} \to \mathbb{R}$ are twice continuously differentiable functions in the neighborhood of x^* where

$$\begin{aligned} &\lim_{x\to x^*} f(x) = \lim_{x\to x^*} g(x) = 0 \text{ or } \lim_{x\to x^*} f(x) = \lim_{x\to x^*} g(x) = +\infty, \text{ then} \\ &\lim_{x\to x^*} \frac{f(x)}{g(x)} = \lim_{x\to x^*} \frac{f'(x)}{g'(x)}. \end{aligned}$$

Example: *Take CRRA utility function $u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$. Then*

$$\lim_{\sigma\to 0} u(c) = \left.\frac{c^{1-\sigma} \ln c}{-1}\right|_{\sigma\to 0} = \ln c.$$

Taylor expansion If function $f(x) : \mathbb{R} \to \mathbb{R}$ is well defined on some range we're interested in, and continuously differentiable till $n+1$-th order, then for a reference point a in this range,

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \ldots + \frac{1}{k!}f^{(k)}(a)(x-a)^k + \ldots.$$

If x is arbitrarily close to a, then the first order Taylor expansion $f(x) = f(a) + f'(a)(x-a)$ is a sufficiently precise approximation.

Example: *$\lim_{x\to 0} \ln(1+x) = x$ is one of the most useful approximation in macroeconomics.*

Implicit function theorem Implicit function is defined through equation $F(x,u) = 0$, $x, u \in \mathbb{R}$ and $u(x) : \mathbb{R} \to \mathbb{R}$. Given $\frac{\partial F}{\partial u} \neq 0$, then

$$\frac{\partial u}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial u}}.$$

Example: *Suppose in equilibrium two macroeconomic variables (c, s) is related by the first order condition $f(c) + g(s) = 0$, and we know that c is a function of s. Then the relation between these two variables can be derived by*

$$f'(c)\frac{\partial c}{\partial s} + g'(s) = 0, \text{ i.e. } \frac{\partial c}{\partial s} = -\frac{g'(s)}{f'(c)}.$$

Log-linearization Suppose that functions $x_1(t)$, $x_2(t)$ are both functions of t. Suppose that function $f(t)$ is the product of them, $f(t) = x_1(t)x_2(t)$, then

$$\frac{\dot{f}(t)}{f(t)} = \frac{\dot{x}_1(t)}{x_1(t)} + \frac{\dot{x}_2(t)}{x_2(t)}.$$

Proof: *$\ln f(t) = \ln x_1(t) + \ln x_2(t)$, and take derivatives of t on both sides.* $\square$

Log-linear approximation Suppose a dynamic system is characterized by

$$f(A_t, B_t, \ldots) = g(Z_t)$$

in which $A_t, B_t, \ldots, Z_t$ are strictly positive variables, and $f(\cdot)$, $g(\cdot)$ may be non-linear in the variables. Also the system has a steady state such that

$f(A, B, \ldots) = g(Z)$. Then

$$\frac{\partial f(A,B,\ldots)}{\partial A}Aa_t + \frac{\partial f(A,B,\ldots)}{\partial B}Bb_t + \ldots = g'(Z)Zz_t, \text{ in which}$$

$$x_t = \ln\left(\frac{X_t}{X}\right) = \ln\left(1 + \frac{X_t - X}{X}\right) \approx \frac{X_t - X}{X},$$

i.e. x_t is approximately the percentage deviation of X_t from the steady state when such deviation is small, capturing the *local* behavior around the steady state.

Proof: *Rewrite the equation using the fact that* $X_t = \exp(\ln X_t)$ *and then take logs on both sides,* $\ln f(\exp(\ln A_t), \exp(\ln B_t), \ldots) = \ln g(\exp(\ln Z_t))$. *Take the first order Taylor approximation around the steady state* $(\ln(A), \ln(B), \ldots, \ln(Z))$

$$\begin{aligned}
&\ln f(A,B,\ldots) + \frac{1}{f(A,B,\ldots)}\left[\frac{\partial f(A,B,\ldots)}{\partial A}A(\ln A_t - \ln A)\right. \\
&\left.+\frac{\partial f(A,B,\ldots)}{\partial B}B(\ln B_t - \ln B) + \ldots\right] \\
&= \ln g(Z) + \frac{1}{g(Z)}\left[g'(Z)Z(\ln Z_t - \ln Z)\right],
\end{aligned}$$

using the definition $x_t = \ln X_t - \ln X$ *and rearrange the equation above to get the result.* □

Leibnitz's rule Suppose that function $f(x) : \mathbb{R} \to \mathbb{R}$ is defined as

$$f(x) = \int_a^{u(x)} g(t)dt$$

in which a is a constant, then

$$f'(x) = g(u(x))u'(x).$$

Proof: *The proof is done in an intuitive way. Calculating* $f(x)$ *means that*

$$f(x) = G(u(x)) - G(a)$$

in which $G(\cdot)$ *is the original function such that* $G'(\cdot) = g(\cdot)$, *and note that* $G(a)$ *is a constant. Then by the chain rule*

$$f'(x) = G'(u(x))u'(x) = g(u(x))u'(x).$$

□

Integration by parts Suppose that $u(x)$ and $v(x)$ are both functions of x and differentiable for $x \in [a, b]$. Then

$$\int_a^b u(x)dv(x) = u(x)v(x)|_a^b - \int_a^b v(x)du(x).$$

Proof: *By the product rule of differentiation*

$$d[u(x)v(x)] = v(x)du(x) + u(x)dv(x).$$

Then integrate both sides on $[a, b]$ *and the result is immediately seen.* $\square$

3 Probability and Statistics

Expectation, variance, and covariance For random variables X with (continuous) probability density function $f(x)$, the expected value of X is

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f(x)dx,$$

the variance of X is

$$\mathrm{var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 .$$

For another random variables Y correlated with X with (continuous) probability density function $g(y)$, the covariance of X and Y is

$$\begin{aligned} \mathrm{cov}[X, Y] &= \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y], \text{ therefore} \\ \mathbb{E}[XY] &= \mathbb{E}[X]\mathbb{E}[Y] + \mathrm{cov}[X, Y]. \end{aligned}$$

Moments of functions of random variables For random variables X with (continuous) probability density function $f(x)$, the moments of its linear transformation $aX + b$ $(a, b \in \mathbb{R})$ are

$$\begin{aligned} \mathbb{E}[aX + b] &= a\mathbb{E}[X] + b, \\ \mathrm{var}[aX + b] &= a^2\mathrm{var}[X]. \end{aligned}$$

For some other random variables Y and Z correlated with X with (continuous) probability density functions $g(y)$, $h(z)$, the moments of linear functions of X, Y, Z ($a, b \in \mathbb{R}$) are

$$\begin{aligned}
\mathrm{cov}[aX + bZ, Y] &= a\mathrm{cov}[X, Y] + b\mathrm{cov}[Z, Y], \\
\mathbb{E}[aX + bY] &= a\mathbb{E}[X] + b\mathbb{E}[Y], \text{ but} \\
\mathrm{var}[aX + bY] &= a^2\mathrm{var}[X] + b^2\mathrm{var}[Y] + 2ab\mathrm{cov}[X, Y].
\end{aligned}$$

Example: *Suppose ϕ and ψ are iid and $\phi \sim N\left(0, \sigma^2\right)$ and $\psi \sim N\left(0, \delta^2\right)$, then what is $cov[a\phi + \psi, \phi]$ ($a \in \mathbb{R}$)?*

$$\mathrm{cov}[a\phi + \psi, \phi] = a\mathrm{cov}[\phi, \phi] + \mathrm{cov}[\psi, \phi].$$

The first term is actually $avar[\phi]$ by the definition of covariance, and the second term is zero since the variables are iid.

Log-normal distribution For a normal distribution $\ln z \sim N\left(\mu, \sigma^2\right)$, the expected value of z is $\mathbb{E}[z] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$, and $\ln \mathbb{E}[z] = \mu + \frac{\sigma^2}{2}$.

Proof: *The probability density distribution function of $x = \ln z$ is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, then*

$$\begin{aligned}
\mathbb{E}[z] &= \int_{-\infty}^{+\infty} e^x f(x) dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x^2 - 2x\mu + \mu^2 - 2\sigma^2 x}{2\sigma^2}\right] dx \\
&= \exp\left(\mu + \frac{\sigma^2}{2}\right) \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu - \sigma^2)^2}{2\sigma^2}\right] dx \\
&= \exp\left(\mu + \frac{\sigma^2}{2}\right).
\end{aligned}$$

An easier approach is to exploit the moment generating function. This is left as an exercise for the readers. □

Linear conditional mean and decision rule For two correlated random variables X and Y, if the conditional mean $\mathbb{E}[X|y]$ is a linear function of y, i.e. the decision rule after observing y is $\mathbb{E}[X|y] = \alpha + \beta y$, then

$$\begin{aligned}
\alpha &= \mathbb{E}[X] - \beta\mathbb{E}[Y], \\
\beta &= \frac{\mathrm{cov}[X, Y]}{\mathrm{var}[Y]}.
\end{aligned}$$

in which $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are unconditional means.

Proof: *Given that the unconditional variance of neither X nor Y is zero,*

$$\mathbb{E}[X|y] = \int_{-\infty}^{+\infty} xf(x|y)dx = \frac{\int_{-\infty}^{+\infty} xf(x,y)dx}{f_Y(y)}, \text{ we get}$$

$$\int_{-\infty}^{+\infty} xf(x,y)dx = f_Y(y)(\alpha + \beta y). \tag{1}$$

Integrate both sides of (1) on y, one can immediately see that $\mathbb{E}[X] = \alpha + \beta\mathbb{E}[Y]$. *Now multiply both sides of (1) by y and integrate both sides of the new equation on y, one can immediately see that* $\mathbb{E}[XY] = \alpha\mathbb{E}[Y] + \beta\mathbb{E}\left[Y^2\right]$. *Then solve these two equations for* α *and* β. □

Bivariate normal distribution and conditional mean For two correlated and *normally distributed* random variables X and Y, the conditional mean $\mathbb{E}[X|y]$ is

$$\mathbb{E}[X|y] = \mathbb{E}[X] + \frac{\mathrm{cov}[X,Y]}{\mathrm{var}[Y]}\left(y - \mathbb{E}[Y]\right)$$

in which $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are unconditional means.

Proof: *Start from the joint distribution of X and Y. — To economize the notations, denote that* $x \sim N(\mu, \sigma^2)$ *and* $y \sim N(\nu, \gamma^2)$

$$f(x,y) = \frac{1}{2\pi\sigma\gamma\sqrt{1-\rho^2}}\exp\left(-\frac{q}{2}\right), \text{ in which}$$

$$q = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu}{\sigma}\right)^2 - 2\rho\left(\frac{x-\mu}{\sigma}\right)\left(\frac{y-\nu}{\gamma}\right) + \left(\frac{y-\nu}{\gamma}\right)^2\right], \text{ and } \mathrm{cov}[X,Y] = \rho\sigma\gamma.$$

$\mathbb{E}[X|y]$ *can be computed from*

$$\mathbb{E}[X|y] = \int_{-\infty}^{+\infty} xf(x|y)dx = \int_{-\infty}^{+\infty} x\frac{f(x,y)}{f_Y(y)}dx,$$

in which $f_Y(y)$ *is the marginal distribution of Y and can be computed from*

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \frac{1}{\gamma\sqrt{2\pi}}\exp\left[-\frac{(y-\nu)^2}{2\gamma^2}\right].$$

Actually in the procedure above one can find that

$$f(x,y) = f_Y(y)\left\{\frac{1}{\sigma\sqrt{1-\rho^2}\sqrt{2\pi}}\exp\left[-\frac{(x-b)^2}{2\sigma^2(1-\rho^2)}\right]\right\}$$

in which $b = \mathbb{E}[X|y] = \mu + \rho\frac{\sigma}{\gamma}(y-\nu) = \mathbb{E}[X] + \frac{cov[X,Y]}{var[Y]}(y - \mathbb{E}[Y])$*, and the result is thus directly seen without further computation.* $\square$