Static optimization

Robert J. Carroll

Last revised: 6 May 2026

---

1 Motivation

The whole project of formal political-economy theory has agents who optimize at its center. A candidate selects a policy position to maximize her vote share. A voter chooses the alternative she most prefers from a menu. A legislator picks a strategy that maximizes her expected payoff in a bargaining game. A regulator chooses the policy that minimizes expected social loss given a budget. A coalition allocates limited resources across a set of districts to maximize the seats it can credibly contest. Each of these is, formally, an optimization problem: a function to be maximized or minimized, possibly subject to a set of constraints, with the agent's choice variable ranging over a feasible set. This handout develops the apparatus for finding the optimum, certifying its existence, and tracking how it shifts when the problem's parameters do.

Three structural questions organize the material. First, when can we identify an optimum by a calculation — by setting a derivative to zero, or by writing down a Lagrangian and reading off the conditions it satisfies? The answer depends on how the constraints are organized: unconstrained problems (§2), equality-constrained problems (§3), and inequality-constrained problems (§4) each get their own treatment, with the gradient, Hessian, and concavity vocabulary supplied by the differentiation and convex-functions handouts that lead in to this one. Second, when does an optimization problem actually *have* a solution? The continuity-and-compactness machinery from the open-and-closed-sets handout is what lets us answer this, and Berge's theorem of the maximum (§5) extends the answer to a parametric setting where we want continuity of the value function and upper-hemicontinuity of the optimum as the parameters vary. Third, how does the optimum change when the parameters do? The envelope theorem (§6) is the structural workhorse of comparative statics — it says that to first order, the value function's response to a parameter change is determined by the parameter's direct effect on the objective, and the maximizer's adjustment can be ignored. Most of the comparative-statics arguments a working political economist actually runs are envelope-theorem applications.

The handout sits as the third in the optimization cluster: differentiation supplied gradients, Hessians, and the implicit function theorem; convex sets and concave functions supplied the local-implies-global maximum theorem under concavity and the differential characterizations that let us check it. The closing handout in the cluster takes up dynamic optimization (the Bellman equation, contraction-mapping arguments, infinite-horizon problems), which builds on the static apparatus established here.

2 Unconstrained optimization

Many of the cleanest optimization problems in formal political-economy theory have no constraints. A voter with quadratic-loss utility maximizes by setting $\nabla u_v = \mathbf{0}$; her optimum is her ideal point. A policymaker free to implement any policy in $\mathbb{R}^n$ chooses by the same calculation. A candidate strategizing without external limits picks $\boldsymbol{p} \in \mathbb{R}^k$ to maximize her vote share unconditionally. Even when constraints are present in principle, the unconstrained problem is usually the structural starting point: identify the unconstrained optimum first, check whether it satisfies the constraints, and only invoke the constrained machinery if it doesn't. The two structural results in this section identify when the calculus condition $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ certifies an optimum — necessarily and (under concavity, leveraging the local-implies-global theorem from the convex-functions handout) sufficiently.

Definition 1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. A point $\boldsymbol{x}^* \in \mathbb{R}^n$ is a *critical point* of f if the gradient vanishes there: $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$.

Theorem 2 (First-order necessary condition). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. If $\boldsymbol{x}^*$ is a local maximum (or local minimum) of f, then $\boldsymbol{x}^*$ is a critical point.*

Proof. A local maximum is a maximum on any sufficiently small neighborhood. If $\nabla f(\boldsymbol{x}^*) \neq \mathbf{0}$, choose any direction $\boldsymbol{v}$ with $\langle \nabla f(\boldsymbol{x}^*), \boldsymbol{v} \rangle > 0$; the function $t \mapsto f(\boldsymbol{x}^* + t\boldsymbol{v})$ has positive derivative at $t = 0$, so f increases along $\boldsymbol{v}$ from $\boldsymbol{x}^*$, contradicting the local-max condition. The mirror argument runs for local minima. ☐

The first-order condition is necessary but not sufficient: critical points include local minima and saddle points, not just local maxima. The second-order condition rules these out using the Hessian's definiteness.

Theorem 3 (Second-order conditions). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable, and let $\boldsymbol{x}^*$ be a critical point.*

- *If the Hessian $H_f(\boldsymbol{x}^*)$ is negative definite, then $\boldsymbol{x}^*$ is a strict local maximum.*
- *If $H_f(\boldsymbol{x}^*)$ is negative semidefinite and $\boldsymbol{x}^*$ is a local maximum, then the second-order condition is consistent (necessity).*
- *If $H_f(\boldsymbol{x}^*)$ is positive definite, $\boldsymbol{x}^*$ is a strict local minimum.*
- *If H_f is indefinite at $\boldsymbol{x}^*$, then $\boldsymbol{x}^*$ is a saddle point.*

Proof sketch. The second-order Taylor expansion gives $f(\boldsymbol{x}^* + \boldsymbol{h}) = f(\boldsymbol{x}^*) + \frac{1}{2}\boldsymbol{h}^\top H_f(\boldsymbol{x}^*)\boldsymbol{h} + o(\|\boldsymbol{h}\|^2)$, since the gradient vanishes. The sign of $\boldsymbol{h}^\top H_f(\boldsymbol{x}^*)\boldsymbol{h}$ for $\boldsymbol{h}$ near $\mathbf{0}$ is determined by the definiteness of $H_f(\boldsymbol{x}^*)$, which controls the local behavior of f around $\boldsymbol{x}^*$. ☐

When the function is concave throughout its domain, the local/global distinction collapses entirely. The convex-functions handout's local-implies-global theorem says that any local maximum of a concave function on a convex set is a global maximum; combined with the FOC-sufficiency corollary (a critical point of a concave function is a global maximum, since the first-order tangent inequality implies $f(\boldsymbol{y}) \leq f(\boldsymbol{x}^*) + \mathbf{0}^\top(\boldsymbol{y} - \boldsymbol{x}^*) = f(\boldsymbol{x}^*)$ for every $\boldsymbol{y}$), this is the workhorse result for unconstrained optimization in PE applications: most quadratic-loss utility functions are strictly concave (their negative-definite Hessians come from the spatial-voting setup), and the FOC pins down the unique global maximum directly.

Example 4 (Spatial voter's utility maximization). A voter with ideal point $\boldsymbol{x}_v \in \mathbb{R}^k$ and quadratic-loss utility $u_v(\boldsymbol{p}) = -(\boldsymbol{p} - \boldsymbol{x}_v)^\top A(\boldsymbol{p} - \boldsymbol{x}_v)$ for symmetric positive-definite A chooses her optimal policy by setting $\nabla u_v = -2A(\boldsymbol{p} - \boldsymbol{x}_v) = \mathbf{0}$, which gives $\boldsymbol{p}^* = \boldsymbol{x}_v$. The Hessian is $-2A$, negative definite by the positive-definiteness of A, so $\boldsymbol{x}_v$ is a strict local (and by concavity, global) maximum. The voter's optimal policy is her ideal point — an unsurprising answer rederived using the unconstrained-optimization apparatus.

3 Equality-constrained optimization

A great deal of political-economy modeling involves resource allocations subject to a fixed budget. A coalition allocating campaign funds across n districts subject to a total-spending constraint. A voter choosing how to split her time across n campaigns subject to a fixed time budget. A regulator allocating enforcement across n regions subject to a personnel constraint. In each case, the choice variable is constrained to lie on a surface in $\mathbb{R}^n$ defined by an equality, and the unconstrained-optimization machinery of the previous section does not apply directly — the gradient at the optimum need not vanish, because the constraint surface bends the feasible set away from the unconstrained gradient direction. The geometric fix — and the structural workhorse of constrained-optimization theory — is the method of Lagrange multipliers: at the optimum, the gradient of the objective is parallel to the gradient of the constraint, with the multiplier giving the proportionality factor.

Theorem 5 (Lagrange multipliers, equality constraint)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable. Suppose $\boldsymbol{x}^*$ is a local maximum of $f(\boldsymbol{x})$ subject to $g(\boldsymbol{x}) = \mathbf{0}$, and the constraint qualification holds: the Jacobian $Dg(\boldsymbol{x}^*)$ has full row rank. Then there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that*

$$\nabla f(\boldsymbol{x}^*) = Dg(\boldsymbol{x}^*)^\top \boldsymbol{\lambda}^*.$$

The conclusion is most naturally written in terms of the *Lagrangian* $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^\top g(\boldsymbol{x})$: the first-order condition is $\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$, which is to say, the unconstrained-optimization first-order condition applied to the Lagrangian. The constraint $g(\boldsymbol{x}) = \mathbf{0}$ recovers the original problem's feasibility condition.

Example 6 (Coalition allocating a fixed budget)**.** A coalition allocates campaign funds $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$ across n districts to maximize total expected seats won, with district-i probability of winning a function $f_i(x_i)$ that is increasing and strictly concave in x_i. The total expected seats are $f(\boldsymbol{x}) = \sum_i f_i(x_i)$, subject to the budget constraint $\sum_i x_i = B$. The Lagrangian is $\mathcal{L}(\boldsymbol{x}, \lambda) = \sum_i f_i(x_i) - \lambda(\sum_i x_i - B)$. The first-order condition gives $f_i'(x_i^*) = \lambda^*$ for every i where $x_i^* > 0$. The structural reading: at the optimum, the marginal expected seat per dollar must be equal across districts; if not, reallocating from a low-marginal district to a high-marginal one would increase the objective without violating the budget. The multiplier λ^* is the common marginal value of an extra dollar of budget — the *shadow price* of the budget constraint, in the formal sense the next paragraph develops.

The Lagrange multiplier λ^* has a substantive economic interpretation: it is the rate at which the optimal value of the objective changes when the constraint is relaxed.[1]

---

[1] The interpretation is sharper than the prose suggests, and worth pinning down. Suppose the constraint is $g(\boldsymbol{x}) = \boldsymbol{b}$ for some right-hand side $\boldsymbol{b}$, and let $V(\boldsymbol{b})$ be the optimal value of the constrained problem as a function of $\boldsymbol{b}$. The envelope-style argument of §6 below gives $\partial V / \partial b_i = \lambda_i^*$: the multiplier is the marginal value of relaxing the ith constraint by one unit. Hence the term *shadow price*: it is the price at which an extra unit of the constraining resource would be valued by the optimizing agent. In the coalition-budget example, λ^* is the marginal expected seat per dollar of budget — the coalition would pay up to λ^* for an extra dollar of campaign funds. The interpretation has bite for political-economy modeling: a binding budget constraint with a high shadow price is exactly the condition under which lobbying for a higher budget is worth it. The interpretation requires the constraint qualification (no-degenerate-Jacobian) to hold; without it, the multiplier may not exist or may not have the marginal-value interpretation. The constraint qualification is generic — it fails only on a measure-zero set of parameter values — but it does fail, and the failure is real (e.g., when constraints are degenerate or redundant). Sundaram (1996, Ch. 5–6) works through the qualification questions in detail; Boyd and Vandenberghe (2004, Ch. 5) treats them in the convex setting where the relevant theory is cleanest.

4 Inequality-constrained optimization

Real political-economy problems usually have inequality constraints rather than (or in addition to) equality constraints. A candidate's spending on persuasion must be *at most* her budget, not exactly equal to it. A regulator's tax rate must be *at least* zero. A coalition's allocation across districts must satisfy non-negativity in each district. The generalization of the Lagrange-multiplier method to inequality constraints is the Karush–Kuhn–Tucker (KKT) conditions, which combine the multiplier setup with a *complementary slackness* condition: a constraint is binding (and so contributes a multiplier) if and only if the multiplier is non-zero.

Theorem 7 (Karush–Kuhn–Tucker)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable, and let* $g : \mathbb{R}^n \to \mathbb{R}^m$ *and* $h : \mathbb{R}^n \to \mathbb{R}^p$ *be continuously differentiable. Suppose* $\boldsymbol{x}^*$ *is a local maximum of* $f(\boldsymbol{x})$ *subject to* $g(\boldsymbol{x}) \leq \mathbf{0}$ *and* $h(\boldsymbol{x}) = \mathbf{0}$*, and a constraint qualification holds (e.g., the Jacobians of binding constraints have full row rank). Then there exist* $\boldsymbol{\mu}^* \in \mathbb{R}^m_+$ *and* $\boldsymbol{\lambda}^* \in \mathbb{R}^p$ *such that:*

- *(Stationarity)* $\nabla f(\boldsymbol{x}^*) = Dg(\boldsymbol{x}^*)^\top \boldsymbol{\mu}^* + Dh(\boldsymbol{x}^*)^\top \boldsymbol{\lambda}^*$*;*
- *(Primal feasibility)* $g(\boldsymbol{x}^*) \leq \mathbf{0}$*,* $h(\boldsymbol{x}^*) = \mathbf{0}$*;*
- *(Dual feasibility)* $\boldsymbol{\mu}^* \geq \mathbf{0}$*;*
- *(Complementary slackness)* $\mu_i^* g_i(\boldsymbol{x}^*) = 0$ *for every* i*.*

The complementary slackness condition is the structural insight. If constraint i is *inactive* at the optimum — $g_i(\boldsymbol{x}^*) < 0$, the constraint is satisfied with slack — then the corresponding multiplier μ_i^* must be zero, so the constraint contributes nothing to the optimum's first-order condition. If the constraint is *active* — $g_i(\boldsymbol{x}^*) = 0$, the constraint binds — then the multiplier may be positive and the constraint contributes its gradient direction to the stationarity condition. So the KKT framework selects, at the optimum, the subset of constraints that are binding and treats them like equality constraints, while ignoring the non-binding ones.

Example 8 (Politician with budget cap and poll-position constraint)**.** A politician maximizes expected vote share $f(\boldsymbol{x})$ where $\boldsymbol{x} = (x_1, x_2)$ is her allocation of effort between persuasion (x_1) and turnout (x_2). She faces two constraints: a budget cap $x_1 + x_2 \leq B$ and a non-negativity constraint $x_1, x_2 \geq 0$. The KKT conditions are: $\nabla f(\boldsymbol{x}^*) = \mu^*(1,1)^\top + \nu_1^*(-1,0)^\top + \nu_2^*(0,-1)^\top$ for non-negative multipliers μ^*, ν_1^*, ν_2^*, with complementary slackness on each constraint. The cases: if the budget binds ($\mu^* > 0$, $\nu_1^* = \nu_2^* = 0$ assuming an interior allocation), the marginal vote shares from persuasion and turnout are equalized at the multiplier value, recovering the marginal-equalization reading of the previous section. If the budget doesn't bind ($\mu^* = 0$), the politician's effort is too low to be limited by funds and her allocation is determined by the marginal-vote-share equalization at zero. The KKT structure walks through which case obtains based on parameter values — a substantive piece of the comparative-statics story.

5 Existence: Weierstrass and Berge's theorem of the maximum

The FOC and SOC identify candidate optima; they do not by themselves guarantee that the optimum *exists*. The empirical question is real. A candidate maximizing vote share over an unbounded policy space might find that the supremum is attained nowhere in particular — if more extreme positions yield ever-higher payoffs without limit, no specific policy is optimal. A continuous-strategy game on an unbounded action set might have no Nash equilibrium because no player has a best response. The existence question is what *Weierstrass's theorem* answers in the static case, and the parametric version — which is what most applied work actually wants, since the analyst typically cares how the optimum moves as the model's parameters move — is what *Berge's theorem of the maximum* answers. Berge is in particular the structural backbone of the existence-of-equilibrium arguments the upcoming game-theory cluster will build on, since it is what guarantees that best-response correspondences are upper-hemicontinuous and non-empty-valued, the input the Kakutani fixed-point theorem requires for an existence proof of Nash equilibrium.

Theorem 9 (Weierstrass)**.** *Let* $f : K \to \mathbb{R}$ *be a continuous function on a non-empty compact set* $K \subseteq \mathbb{R}^n$*. Then* f *attains its maximum and minimum on* K*.*

The proof is the classical compactness-plus-continuity argument from §5 of the open-and-closed-sets handout: a continuous image of a compact set is compact, and a compact subset of $\mathbb{R}$ has a maximum and a minimum. The PE reading: if the strategy space is compact (e.g., closed and bounded, by Heine–Borel) and the objective is continuous, an optimum exists. This is the structural justification for compactness assumptions on strategy spaces in game theory and bargaining theory.

The parametric version asks: when the optimization problem depends on a parameter $\boldsymbol{\theta}$, how do the optimal value and the set of optimizers depend on $\boldsymbol{\theta}$? Berge's theorem is the structural answer.

Definition 10. Let $\Gamma : \Theta \rightrightarrows \mathbb{R}^n$ be a correspondence (a set-valued map: each $\boldsymbol{\theta}$ has $\Gamma(\boldsymbol{\theta}) \subseteq \mathbb{R}^n$). Γ is *upper-hemicontinuous* (uhc) at $\boldsymbol{\theta}_0$ if, for every open $U \supseteq \Gamma(\boldsymbol{\theta}_0)$, there is a neighborhood V of $\boldsymbol{\theta}_0$ such that $\Gamma(\boldsymbol{\theta}) \subseteq U$ for every $\boldsymbol{\theta} \in V$. Γ is *lower-hemicontinuous* (lhc) at $\boldsymbol{\theta}_0$ if, for every open U with $\Gamma(\boldsymbol{\theta}_0) \cap U \neq \emptyset$, there is a neighborhood V of $\boldsymbol{\theta}_0$ such that $\Gamma(\boldsymbol{\theta}) \cap U \neq \emptyset$ for every $\boldsymbol{\theta} \in V$. Γ is *continuous* if it is both uhc and lhc.

The intuition: uhc says the correspondence does not "explode upward" — Γ's graph cannot suddenly include new far-away points as the parameter is perturbed. Lhc says it does not "collapse" — Γ cannot suddenly drop a previously-included point. For a single-valued correspondence (a function), both notions reduce to ordinary continuity.

Theorem 11 (Berge's theorem of the maximum)**.** *Let* $f : X \times \Theta \to \mathbb{R}$ *be continuous, and let* $\Gamma : \Theta \rightrightarrows X$ *be a continuous correspondence with non-empty compact values. Define*

$$V(\boldsymbol{\theta}) = \max_{\boldsymbol{x} \in \Gamma(\boldsymbol{\theta})} f(\boldsymbol{x}, \boldsymbol{\theta}), \qquad \Gamma^*(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{x} \in \Gamma(\boldsymbol{\theta})} f(\boldsymbol{x}, \boldsymbol{\theta}).$$

Then $V : \Theta \to \mathbb{R}$ *is continuous, and* $\Gamma^* : \Theta \rightrightarrows X$ *is upper-hemicontinuous with non-empty compact values.*

The proof combines Weierstrass on each $\Gamma(\boldsymbol{\theta})$ (giving non-emptiness and compactness of Γ^*) with the continuity of f and Γ to get continuity of V and uhc of Γ^*. Sundaram (1996, Ch. 9) works through the proof; Aliprantis and Border (2006, Ch. 17) gives the general topological-space version.

The political-economy reading: in any game where each player's best-response correspondence depends continuously on the others' strategies (continuous payoff function on a continuous compact strategy correspondence), Berge's theorem ensures the best-response correspondence itself is upper-hemicontinuous and non-empty-valued. This is exactly the input the Kakutani fixed-point theorem (game-theory cluster) requires for an existence proof of Nash equilibrium. Berge is in this sense the structural backbone of equilibrium-existence arguments in formal political theory.

6 The envelope theorem and comparative statics

Comparative statics — how does the optimal choice or the optimal value change when a parameter shifts? — is the central methodological move in formal political-economy theory. A change in the cost of campaigning shifts the candidate's optimal effort allocation. A change in the discount factor shifts the legislator's optimal proposal. A change in the prior over a state shifts the regulator's optimal policy. In every case, the analyst wants to know how the value function $V(\boldsymbol{\theta}) = \max_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{\theta})$ depends on $\boldsymbol{\theta}$. The naive approach is to differentiate: $\partial V/\partial\theta_i = \partial f/\partial\theta_i$ at the optimum, plus a correction term capturing how the maximizer $\boldsymbol{x}^*(\boldsymbol{\theta})$ moves with $\boldsymbol{\theta}$. The *envelope theorem* says the correction term vanishes: to first order, the maximizer's adjustment is irrelevant, and the only thing that matters is how $\boldsymbol{\theta}$ shifts f directly at the unchanged optimum.

The structural reason is the FOC. Around the maximizer, small perturbations of $\boldsymbol{x}$ do not change f to first order (because $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}) = \mathbf{0}$), so the maximizer's movement contributes nothing first-order. Only the parameter's direct effect survives.

Theorem 12 (Envelope theorem, unconstrained case)**.** *Let $f : X \times \Theta \to \mathbb{R}$ be continuously differentiable on an open set, and suppose $\boldsymbol{x}^*(\boldsymbol{\theta}) \in \operatorname{int} X$ is a continuously differentiable interior maximizer of $f(\cdot, \boldsymbol{\theta})$ for $\boldsymbol{\theta}$ in an open set. Define $V(\boldsymbol{\theta}) = f(\boldsymbol{x}^*(\boldsymbol{\theta}), \boldsymbol{\theta})$. Then*

$$\frac{\partial V}{\partial \theta_i}(\boldsymbol{\theta}) = \frac{\partial f}{\partial \theta_i}(\boldsymbol{x}^*(\boldsymbol{\theta}), \boldsymbol{\theta}).$$

Proof. Differentiating $V(\boldsymbol{\theta}) = f(\boldsymbol{x}^*(\boldsymbol{\theta}), \boldsymbol{\theta})$ via the chain rule:

$$\frac{\partial V}{\partial \theta_i} = \nabla_{\boldsymbol{x}} f \cdot \frac{\partial \boldsymbol{x}^*}{\partial \theta_i} + \frac{\partial f}{\partial \theta_i} = \mathbf{0} + \frac{\partial f}{\partial \theta_i},$$

using $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}^*(\boldsymbol{\theta}), \boldsymbol{\theta}) = \mathbf{0}$ (the FOC). ☐

The constrained version is the same idea applied to the Lagrangian: the value function's derivative with respect to $\boldsymbol{\theta}$ equals the partial derivative of the Lagrangian with respect to $\boldsymbol{\theta}$, evaluated at the optimum.

Theorem 13 (Envelope theorem, constrained case)**.** *Let f, g be continuously differentiable, and let $V(\boldsymbol{\theta}) = \max_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{\theta})$ subject to $g(\boldsymbol{x}, \boldsymbol{\theta}) = \mathbf{0}$. Suppose the maximizer $\boldsymbol{x}^*(\boldsymbol{\theta})$ and multiplier $\boldsymbol{\lambda}^*(\boldsymbol{\theta})$ vary continuously differentiably with $\boldsymbol{\theta}$, and the constraint qualification holds. Define the Lagrangian $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\theta}) = f(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{\lambda}^\top g(\boldsymbol{x}, \boldsymbol{\theta})$. Then*

$$\frac{\partial V}{\partial \theta_i}(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \theta_i}(\boldsymbol{x}^*(\boldsymbol{\theta}), \boldsymbol{\lambda}^*(\boldsymbol{\theta}), \boldsymbol{\theta}).$$

The constrained version generalizes to the inequality-constrained KKT case (with multipliers $\boldsymbol{\mu}^*$ on inequality constraints) by the same envelope argument applied to the Lagrangian-with-KKT.[2]

---

[2] The envelope theorem is the structural workhorse of comparative statics, and a substantial literature uses it to derive sharp predictions about how political-economy systems respond to parameter changes. A useful generalization, due to Milgrom and Segal (2002), weakens the differentiability hypotheses considerably: an envelope-style identity holds for value functions even when the maximizer is not differentiable in the parameter, as long as the objective is differentiable in the parameter pointwise. The Milgrom–Segal version is what is invoked in mechanism-design and auction-theory applications where the maximizer's behavior may have kinks. A different generalization, the theory of

Example 14 (Comparative statics: how does optimal effort respond to a cost shift?)**.** Continuing Example 8: suppose the politician's expected vote-share function is $f(\boldsymbol{x}, \theta) = -\frac{1}{2}(x_1 - \theta)^2 - \frac{1}{2}(x_2 - \theta)^2$, where θ is a campaign-effectiveness parameter. The unconstrained optimum is $x_1^* = x_2^* = \theta$; the value function is $V(\theta) = 0$. By the envelope theorem, $\partial V/\partial\theta = \partial f/\partial\theta = (x_1 - \theta) + (x_2 - \theta)$ evaluated at $\boldsymbol{x}^*$, which is zero at the optimum — consistent with the constant value function. Now suppose the budget constraint $x_1 + x_2 \leq B$ binds at the optimum ($B < 2\theta$). The constrained optimum is $x_1^* = x_2^* = B/2$ with multiplier $\mu^* = \theta - B/2$. The value function is $V(\theta, B) = -2(B/2 - \theta)^2 = -(B - 2\theta)^2/2$. The envelope theorem gives $\partial V/\partial B = -\mu^*$ (the shadow price of the budget) and $\partial V/\partial\theta = 2(\theta - B/2) = 2\mu^*$, both checkable by direct differentiation. The envelope-theorem reading: increasing the campaign-effectiveness parameter θ raises the politician's value by $2\mu^*$ per unit, with the proportionality factor being the shadow price of her binding budget constraint — a politically substantive prediction obtained without re-solving the optimization problem.

7 What's next

The next handout takes up dynamic optimization: finite- and infinite-horizon problems, the Bellman equation, contraction-mapping arguments, and the envelope condition for value functions. The static apparatus of this handout is the structural starting point; dynamic optimization adds the recursive structure that lets the same machinery handle problems unfolding over time. The political-economy applications — voter learning over multiple elections, dynamic legislative bargaining, infinite-horizon policymaking under discounting — are central to a substantial literature in formal theory.

Beyond the cluster, the optimization machinery here is the structural prerequisite for game theory (best-response correspondences via Berge, equilibrium existence via fixed-point arguments, comparative statics via envelope) and for the constrained-allocation literature in mechanism design. The next applied cluster — almost certainly game theory — will lean heavily on this material.

For graduate-level treatments at this handout's level of abstraction, Sundaram (1996) is the standard PE-flavored reference and works at exactly this level. Mas-Colell, Whinston, and Green (1995, Ch. 3, M.K) treats the same material in a microeconomic-theory context. Boyd and Vandenberghe (2004) is the canonical convex-optimization reference (freely available online), with a sharper treatment of duality and constraint qualification. Milgrom and Segal (2002) is the source for the modern envelope theorem; Topkis (1998) for monotone comparative statics.

8 Exercises

Exercise 15. Continuing Example 4: take $k = 2$, $\boldsymbol{x}_v = (1, 2)$, and $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$. (a) Write out $u_v(\boldsymbol{p})$ explicitly. (b) Verify that A is positive definite (check its leading principal minors via Sylvester's

---

monotone comparative statics due to Milgrom and Shannon (1994) and Topkis (1998), drops differentiability entirely and gives sharp results on how the maximizer's *ordinal* response to parameter changes is governed by a property of the objective called *supermodularity.* Supermodular games (those where every player's marginal payoff to her own strategy is increasing in others' strategies) have especially well-behaved equilibria, and the comparative-statics tools from monotone-comparative-statics theory transfer directly. The political-economy applications — regulatory tournaments with strategic complementarity, policy diffusion across jurisdictions, comparative statics in legislative bargaining — are substantial. Topkis (1998) is the canonical reference; Sundaram (1996, Ch. 10) works through the basics in the context of optimization theory.

criterion from the previous handout). (c) Compute the FOC and SOC at $\boldsymbol{p}^* = \boldsymbol{x}_v$ and verify the second-order condition is satisfied.

Exercise 16. Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is strictly concave and has a critical point, that critical point is the unique global maximum. (Hint: suppose $\boldsymbol{x}^*$ is a critical point and $\boldsymbol{y} \neq \boldsymbol{x}^*$. Use strict concavity to show $f((\boldsymbol{x}^* + \boldsymbol{y})/2) > (f(\boldsymbol{x}^*) + f(\boldsymbol{y}))/2$, and combine with the FOC to derive $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$.)

Exercise 17. *Equality-constrained allocation, two districts.* A coalition allocates a budget B across two districts to maximize $f_1(x_1) + f_2(x_2)$ subject to $x_1 + x_2 = B$, with $f_i(x) = \log(1 + x)$. (a) Compute the FOC via the Lagrangian. (b) Solve for x_1^*, x_2^* and the multiplier λ^*. (c) Interpret λ^* as a shadow price: how much would the coalition pay for an extra dollar of budget?

Exercise 18. *KKT and binding constraints.* A regulator maximizes $f(x_1, x_2) = -x_1^2 - 2x_2^2 + 4x_1 + 8x_2$ subject to $x_1, x_2 \geq 0$ and $x_1 + x_2 \leq 5$. (a) Compute the unconstrained maximum and check whether it satisfies the constraints. (b) Apply the KKT conditions and identify which constraints bind at the optimum. (c) Compute the optimum and the multipliers.

Exercise 19. *Failure of Weierstrass without compactness.* Give two examples: (a) a continuous function on a non-compact subset of $\mathbb{R}$ that has no maximum; (b) a discontinuous function on a compact subset of $\mathbb{R}$ that has no maximum. Explain in one sentence which hypothesis of Weierstrass's theorem fails in each case.

Exercise 20. *Berge applied to best responses.* In a two-player game, each player chooses an action from a compact strategy set $S_i \subseteq \mathbb{R}^k$, with payoff $u_i(s_i, s_{-i})$ continuous in both arguments. Define player i's best-response correspondence $B_i(s_{-i}) = \arg\max_{s_i \in S_i} u_i(s_i, s_{-i})$. (a) Use Berge's theorem to argue that B_i is upper-hemicontinuous and non-empty-valued. (b) Why does this matter for proving the existence of Nash equilibrium?

Exercise 21. Verify the envelope theorem on $f(x, \theta) = -\frac{1}{2}(x - \theta)^2 + \theta x$. (a) Compute $x^*(\theta)$ from the FOC and the value function $V(\theta) = f(x^*(\theta), \theta)$. (b) Differentiate V directly to find $\partial V/\partial\theta$. (c) Verify it equals $\partial f/\partial\theta$ evaluated at $x^*(\theta)$.

Exercise 22. *Envelope theorem, constrained version.* A coalition allocates a budget B across two districts as in Exercise 17. (a) Compute the value function $V(B)$. (b) Compute $V'(B)$ directly. (c) Verify it equals the multiplier $\lambda^*(B)$ at the optimum.

Exercise 23. Give an example of a correspondence $\Gamma : \mathbb{R} \rightrightarrows \mathbb{R}$ that is upper-hemicontinuous but not lower-hemicontinuous, and another that is lhc but not uhc. (Hint: think about correspondences whose graph has a vertical jump — the direction of the jump determines which hemicontinuity fails.)

Exercise 24. *Comparative statics in legislative bargaining.* A legislator chooses a proposal $x \in \mathbb{R}$ to maximize $u(x, \theta) = -(x - \theta)^2 - cx^2$, where θ is the legislator's ideal point and c is the cost of extreme proposals (a parameter capturing voter backlash, party discipline, etc.). (a) Compute the optimal proposal $x^*(\theta, c)$. (b) Compute the value function $V(\theta, c)$. (c) Use the envelope theorem to compute $\partial V/\partial c$ and interpret: how does an increase in the backlash parameter affect the legislator's payoff at her own optimum?

References

Aliprantis, Charalambos D. and Kim C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide.* 3rd ed. Berlin: Springer.

Boyd, Stephen and Lieven Vandenberghe (2004). *Convex Optimization*. Cambridge: Cambridge University Press.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995). *Microeconomic Theory*. New York: Oxford University Press.

Milgrom, Paul and Ilya Segal (2002). “Envelope Theorems for Arbitrary Choice Sets”. In: *Econometrica* 70.2, pp. 583–601.

Milgrom, Paul and Chris Shannon (1994). “Monotone Comparative Statics”. In: *Econometrica* 62.1, pp. 157–180.

Sundaram, Rangarajan K. (1996). *A First Course in Optimization Theory*. Cambridge: Cambridge University Press.

Topkis, Donald M. (1998). *Supermodularity and Complementarity*. Princeton: Princeton University Press.