

Optimización y Control

Ignacio Delgado, Ignacio Gómez,
José M. Perales, José M. Vega

Escuela de Ingeniería Aeronáutica y del Espacio

Programa: **Grado en Ingeniería Aeronáutica**

Curso 2017-2018

Tema #3: Extremos locales con restricciones

Index

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 Multiplicadores de Lagrange y condiciones KKT
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 Cálculo variacional en dimensión infinita
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

OPTIMIZACIÓN LOCAL CON RESTRICCIONES

Se trata ahora de abordar, mediante **métodos locales**, el cálculo de mínimos locales con restricciones, es decir, calcular:

$$\min F(u_1, \dots, u_N)$$

en un **espacio de diseño** definido por

$$r_k(u_1, \dots, u_N) = 0, \quad \text{para } k = 1, \dots, m, \text{ con } m < N,$$

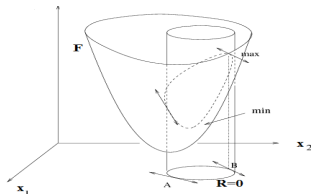
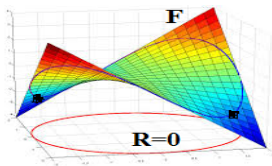
$$r_k(u_1, \dots, u_N) \geq 0, \quad \text{para } k = m + 1, \dots, m + n, \text{ con } m + n < N.$$

- Las primeras m restricciones son **restricciones estrictas**, y deben definir una variedad suave de dimensión $N - m$.
- Las restantes $n - m$ restricciones son **restricciones unilaterales** que definen la frontera del dominio. Por ejemplo, en dimensión 3, si el dominio es un poliedro, las restricciones unilaterales definen las caras del poliedro, cuyo número puede ser cualquiera.

RESTRICCIONES ERICTAS

Cada restricción estricta define una **hipersuperficie** del espacio paramétrico (de dimensión $N - 1$). Para asegurar que el dominio admisible es una **variedad suave**, se hacen las siguientes **hipótesis generales** respecto de las restricciones estrictas:

- Las restricciones son compatibles. Por ejemplo, no debe permitirse que dos de las restricciones estrictas definan hipersuperficies paralelas.
- Las **restricciones estrictas** son tales que **los vectores $\nabla r_1, \dots, \nabla r_m$ son linealmente independientes** entre sí. Por ejemplo, no se admite que dos de las hipersuperficies sean tangentes en el punto.



RESTRICCIONES UNILATERALES

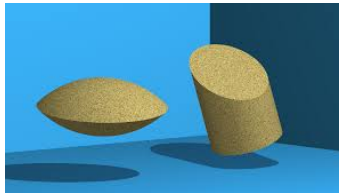
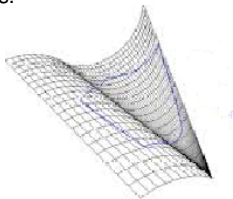
En cada punto de la frontera, una al menos de las restricciones unilaterales es estricta (es decir, verifica la igualdad, en vez de la desigualdad estricta).

Esas restricciones se conocen como **restricciones activas en ese punto**; el resto son restricciones inactivas.

Por ejemplo, si la frontera es un poliedro, en los puntos de las caras, aristas y vértices hay una, dos y tres restricciones activas, respectivamente.

Para asegurar que la frontera es suave a trozos, se hace la siguiente **hipótesis** para las restricciones unilaterales:

Las **restricciones estrictas activas** en cada punto de la frontera son tales que **los correspondientes vectores** ∇r_j son (i) **linealmente independientes** entre sí y (ii) linealmente independientes de los vectores $\nabla r_1, \dots, \nabla r_m$ asociados a las restricciones estrictas.



COMENTARIOS PREVIOS

Históricamente, la optimización con restricciones se empezó tratando mediante métodos de penalización y funciones barrera, que reducen el problema a otro sin restricciones. Estos métodos se basan en ideas sencillas pero no son muy eficaces computacionalmente y requieren cierto calibrado.

Más recientemente se han desarrollado métodos basados en ideas teóricas bien conocidas desde el siglo XIX pero que requieren una implementación computacional más sofisticada. Estos métodos constituyen el estado del arte en la actualidad.

Los mínimos locales del problema anterior pueden clasificarse en dos tipos:

- **Interiores al espacio de diseño:** en puntos que satisfacen estrictamente todas las restricciones unilaterales. Los posibles extremos se calculan mediante el **método de multiplicadores de Lagrange**.
- Pertenecientes a la **frontera del espacio de diseño:** en puntos que satisfacen de modo no estricto alguna(s) de las restricciones unilaterales (restricciones activas). Los posibles extremos se calculan también mediante los multiplicadores de Lagrange, convirtiendo en estrictas las restricciones unilaterales activas. Para asegurar que pueden, en efecto, ser mínimos, la función objetivo debe ser creciente hacia el interior del dominio, lo cual se asegura mediante las **condiciones de Karush-Khun-Tucker (KKT)**.

MÉTODOS DE PENALIZACIÓN

- La idea de los métodos de penalización es modificar la función objetivo de modo que ésta crezca mucho cuando no se verifican las restricciones; por tanto, si no se verifican, 'se penaliza' la función objetivo.
- Para ello, por ejemplo, se añaden nuevos sumandos a la función objetivo, que son proporcionales al cuadrado de las restricciones (**funciones de penalización**) y **parámetros de penalización**, que se hace tender a infinito.
- La nueva función objetivo, sin restricciones, se optimiza con los métodos del tema anterior (que, ¡ojo! pueden dar problemas para valores grandes de los parámetros de penalización).
- Los distintos métodos difieren en la forma de las funciones de penalización.
- Son métodos que requieren calibrado, en cada caso particular.

MÉTODO DE COURANT

Para resolver el problema

$\min F(u_1, \dots, u_N)$, con

$r_k(u_1, \dots, u_N) = 0$, para $k = 1, \dots, m$,

$r_k(u_1, \dots, u_N) \geq 0$, para $k = m + 1, \dots, n$,

se considera la función objetivo

$$F(u_1, \dots, u_N) + \sum_{k=1}^m \sigma_k r_k(u_1, \dots, u_N)^2 + \sum_{k=m+1}^n \sigma_k \min\{r_k(u_1, \dots, u_N), 0\}^2.$$

Se comienza resolviendo el problema de optimización modificado con valores iniciales de los parámetros de penalización (p.e., $\sigma_1 = \dots = \sigma_N = 1$) y se van aumentando en cada paso (p.e., multiplicándolos por 10) hasta conseguir la convergencia.

OTROS MÉTODOS DE PENALIZACIÓN

- Modificando la función de penalización, que es cuadrática en el método de Courant, se obtienen otros métodos que pueden ser más eficaces en situaciones concretas.
- Si se utilizan métodos de tipo gradiente para minimizar la función objetivo penalizada, los problemas resultantes suelen estar mal condicionados cuando el parámetro de penalización es grande y el óptimo es cercano a la frontera del dominio admisible.

FUNCIONES BARRERA

Se trata de un método similar a los métodos de penalización para tratar, solamente, las restricciones unilaterales. Éstas se sustituyen por añadir a la función objetivo una función barrera, que tiende a infinito cuando el punto se acerca a la frontera del dominio admisible.

Las funciones barrera más utilizadas son la función inversa y la función logarítmica, que darían lugar a funciones objetivo del tipo

$$F(u_1, \dots, u_N) + \sum_{k=m+1}^n \sigma_k / r_k(u_1, \dots, u_N),$$

y

$$F(u_1, \dots, u_N) - \sum_{k=m+1}^n \sigma_k \log[r_k(u_1, \dots, u_N)].$$

En ambos casos se minimiza la función objetivo modificada para valores decrecientes del parámetro $\sigma_k > 0$.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 **Multiplicadores de Lagrange y condiciones KKT**
 - **Multiplicadores de Lagrange**
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 Cálculo variacional en dimensión infinita
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

ENFOQUE VARIACIONAL

Considerando sólo las **restricciones estrictas** (es decir, solamente mínimos locales en el interior del dominio admisible), se busca

$$\min F(u_1, \dots, u_N), \text{ con } r_k(u_1, \dots, u_N) = 0, \quad \text{para } k = 1, \dots, m.$$

- Utilizando un **enfoque variacional**, se consideran curvas del espacio paramétrico que verifiquen las restricciones, $\mathbf{u} = \mathbf{u}_0 + \tilde{\mathbf{u}}(\alpha)$, y se impone que la función objetivo tenga un mínimo a lo largo de las curvas en el punto. La condición es entonces $dF(\mathbf{u}_0 + \tilde{\mathbf{u}}(0))/d\alpha = 0$ para toda curva, es decir

$$\nabla F(\mathbf{u}_0)^\top \tilde{\mathbf{u}}'_0 = 0 \text{ con } \tilde{\mathbf{u}}'_0 = \tilde{\mathbf{u}}'(0).$$

- Por otro lado, las curvas deben verificar las restricciones, es decir, $r_k(\mathbf{u}_0 + \tilde{\mathbf{u}}(\alpha)) = 0$ para todo k y para todo α . Derivando respecto de α y particularizando en $\alpha = 0$, se obtiene

$$\nabla r_k(\mathbf{u}_0)^\top \tilde{\mathbf{u}}'_0 = 0 \text{ para } k = 1, \dots, m.$$

MULTIPLICADORES DE LAGRANGE

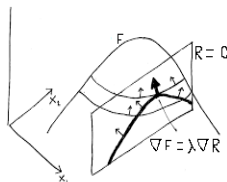
- Resumiendo lo anterior, se tiene

$$\nabla F(\mathbf{u}_0)^\top \tilde{\mathbf{u}}'_0 = 0, \text{ si } \nabla r_k(\mathbf{u}_0)^\top \tilde{\mathbf{u}}'_0 = 0 \text{ para } k = 1, \dots, m.$$

- Es decir, ahora $\nabla F(\mathbf{u}_0)$ no tiene porqué ser ortogonal a todos los vectores de \mathbb{R}^N (lo cual hubiera implicado que $\nabla F(\mathbf{u}_0) = \mathbf{0}$), sino sólo a algunos de ellos.
- $\nabla F(\mathbf{u}_0)$ debe ser ortogonal a todos los vectores que sean ortogonales, simultáneamente, a $\nabla r_1, \dots, \nabla r_m$. Por tanto, debe ser linealmente dependiente de esos vectores, es decir,

$$\nabla F(\mathbf{u}_0) = \lambda_1 \nabla r_1(\mathbf{u}_0) + \dots + \lambda_m \nabla r_m(\mathbf{u}_0).$$

- Los escalares $\lambda_1, \dots, \lambda_m$ se conocen como **multiplicadores de Lagrange**.



SISTEMA DE LAGRANGE Y FUNCIÓN LAGRANGIANA

- Por tanto, los posibles extremos locales del problema con restricciones deben verificar el siguiente **sistema de Lagrange**:

$$\begin{aligned}\nabla F(\mathbf{u}_0) &= \lambda_1 \nabla r_1(\mathbf{u}_0) + \dots + \lambda_m \nabla r_m(\mathbf{u}_0), \\ r_k(\mathbf{u}_0) &= 0, \quad \text{para } k = 1, \dots, m.\end{aligned}$$

- Se tiene así una ecuación vectorial de orden N (que equivale a N ecuaciones escalares) y m ecuaciones escalares, es decir, $N + m$ ecuaciones. Y hay que calcular una incógnita vectorial ($\mathbf{u}_0 \in \mathbb{R}^N$) y m incógnitas escalares, es decir, $N + m$ incógnitas.
- Es decir, se tienen tantas ecuaciones como incógnitas.
- El sistema es lineal en los multiplicadores de Lagrange, pero, en general, es no lineal en su dependencia de \mathbf{u}_0 .
- El sistema de Lagrange proporciona los puntos estacionarios de la **función de Lagrange**, o **Lagrangiana** (dependiente de $N + m$ variables)

$$L(u_1, \dots, u_N; \lambda_1, \dots, \lambda_m) = F(u_1, \dots, u_N) - \sum_{k=1}^m \lambda_k r_k(u_1, \dots, u_N).$$

JACOBIANO DEL SISTEMA DE LAGRANGE

- La matriz jacobiana del sistema de Langrange (dos transparencias atrás) puede escribirse por bloques como

$$\mathbf{J} = \begin{bmatrix} \mathbf{H}^L & -\nabla \mathbf{r}^\top \\ -\nabla \mathbf{r} & \mathbf{O} \end{bmatrix},$$

donde \mathbf{H}^L es la matriz hessiana de la función de Lagrange respecto de las variables de diseño, es decir, $\mathbf{H}_{ij}^L = \partial_{u^i u^j} (F - \sum_{k=1}^m \lambda_k r_k)$; $\nabla \mathbf{r}$ es la matriz jacobiana del vector de restricciones $\mathbf{r} = (r_1, \dots, r_m)$; y \mathbf{O} es la matriz nula de orden $m \times m$.

- En condiciones genéricas, la matriz \mathbf{J} es no singular.
- La matriz \mathbf{H}^L es simétrica y, por tanto, la matriz \mathbf{J} es también simétrica. Pero \mathbf{J} no es, en general, definida positiva, sino indefinida.
- \mathbf{J} es también la matriz hessiana de la función Lagrangiana, considerando ésta como función de $N + m$ variables. Pero, como \mathbf{J} es indefinida, los mínimos de F no son mínimos de la Lagrangiana (considerando ésta como función de $N + m$ variables), sino puntos de ensilladura.

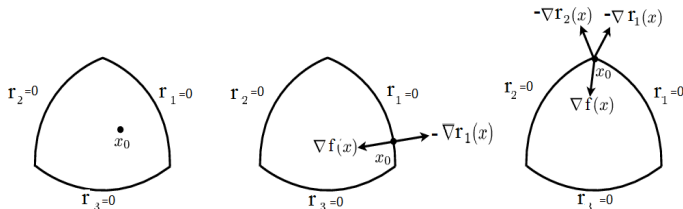
MATRIZ HESSIANA RESTRINGIDA

En los puntos donde se verifican las restricciones, $L = F$. Es decir, la Lagrangiana coincide con la restricción de la función objetivo a la variedad no lineal definida por las restricciones, \hat{F} , y se tiene:

- El vector $\nabla_u L$ puede verse como el vector gradiente de \hat{F} .
- La matriz hessiana $\nabla_{uu}^2 L$ puede verse como la matriz hessiana de \hat{F} .

Se tienen así las siguientes condiciones para la existencia de mínimo:

- Primera condición necesaria:** $\nabla_u L = 0$ (es el sistema de Lagrange).
- Segunda condición necesaria:** La matriz hessiana restringida, $\nabla_{uu}^2 L$, debe ser semidefinida positiva, es decir, todos sus autovalores deben ser ≥ 0 .
- Condición suficiente:** La matriz hessiana restringida, $\nabla_{uu}^2 L$, es definida positiva, es decir, todos sus autovalores son > 0 .



MÉTODO DE NEWTON Y PROGRAMACIÓN CUADRÁTICA (I)

El sistema de Lagrange puede resolverse iterativamente. Dos **formulaciones equivalentes** para cada paso de la iteración:

- **Método de Newton aplicado al sistema de Lagrange**, $\partial_{\mathbf{u}} f = \lambda^\top \partial_{\mathbf{u}} \mathbf{r}$, $\mathbf{r} = \mathbf{0}$: proporciona $(\mathbf{u}_{k+1}, \lambda_{k+1})$ como la única solución del sistema

$$\begin{aligned} [\partial_{\mathbf{u}\mathbf{u}} f - \lambda^\top \partial_{\mathbf{u}\mathbf{u}} \mathbf{r}]_k (\mathbf{u} - \mathbf{u}_k) - (\lambda - \lambda_k)^\top [\partial_{\mathbf{u}} \mathbf{r}]_k &= -[\partial_{\mathbf{u}} f - \lambda^\top \partial_{\mathbf{u}} \mathbf{r}]_k, \\ [\partial_{\mathbf{u}} \mathbf{r}]_k (\mathbf{u} - \mathbf{u}_k) &= -\mathbf{r}_k, \end{aligned}$$

que se obtiene linealizando el sistema de Lagrange y tiene solución única si $[\partial_{\mathbf{u}\mathbf{u}} f - \lambda^\top \partial_{\mathbf{u}\mathbf{u}} \mathbf{r}]_k$ es no singular y $[\partial_{\mathbf{u}} \mathbf{r}]_k$ tiene rango máximo.

- **Programación cuadrática**: $(\mathbf{u}_{k+1}, \lambda_{k+1})$ minimiza el siguiente problema cuadrático con restricciones lineales

$$\begin{aligned} \min_{\mathbf{u}} f_k + [\partial_{\mathbf{u}} f - \lambda^\top \partial_{\mathbf{u}} \mathbf{r}]_k (\mathbf{u} - \mathbf{u}_k) + \frac{1}{2} (\mathbf{u} - \mathbf{u}_k)^\top [\partial_{\mathbf{u}\mathbf{u}} f - \lambda^\top \partial_{\mathbf{u}\mathbf{u}} \mathbf{r}]_k (\mathbf{u} - \mathbf{u}_k) \\ [\partial_{\mathbf{u}} \mathbf{r}]_k (\mathbf{u} - \mathbf{u}_k) = -\mathbf{r}_k, \end{aligned}$$

que se obtiene reteniendo términos de segundo orden (respecto de \mathbf{u}) en la función lagrangiana y linealizando la restricción.

MÉTODO DE NEWTON Y PROGRAMACIÓN CUADRÁTICA (II)

Teniendo en cuenta las conclusiones en optimización sin restricciones:

- El **método de Newton** puede implementarse combinando (el propio método de Newton o el BFGS) con **regiones de confianza**; proporciona **puntos estacionarios** (no necesariamente mínimos) del problema con restricciones. Pero monitorizando los valores de la función objetivo puede asegurarse la convergencia sólo a **mínimos**.
- La **programación cuadrática** puede combinarse con **métodos de descenso**, preferiblemente con descenso basado en BFGS o con gradiente conjugado. Para mejorar las aproximaciones lineales de las restricciones, éstas pueden re-imponerse penalizando la función objetivo cuadrática.

Al igual que en optimización sin restricciones, se obtienen **métodos convergentes para condiciones iniciales genéricas**.

Las aplicaciones MatLab se harán directamente para el caso general, que se considera en el siguiente epígrafe.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 **Multiplicadores de Lagrange y condiciones KKT**
 - Multiplicadores de Lagrange
 - **Condiciones de Karush-Khun-Tucker**
 - Optimización mediante la herramienta MatLab
- 4 Cálculo variacional en dimensión infinita
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

RESTRICCIONES ERICTAS Y UNILATERALES

Para el problema general hay que **resolver el sistema de Lagrange considerando las restricciones activas**, que son las estrictas en los puntos interiores y algunas unilaterales en la frontera. Pero:

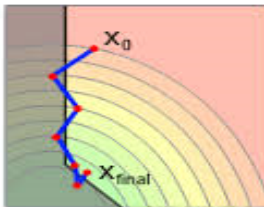
- La opción obvia (fuerza bruta), es decir, **minimizar por separado** en el interior y en cada parte suave de la frontera y tomar el menor de los mínimos así obtenidos **es una estrategia inviable** (salvo cuando hay muy pocas restricciones unilaterales): conduce a una cantidad excesiva de combinaciones de restricciones estrictas a considerar.
- No todos los mínimos de estos problemas son válidos: hay que tener en cuenta el **comportamiento hacia el interior** (condiciones KKT).

Se han desarrollado métodos, conocidos genéricamente como **métodos de conjunto activo (active set)**: en cada interacción, identifican las restricciones activas (que pueden variar de una iteración a otra) y resuelven el correspondiente problema de Lagrange, monitorizando las restricciones inactivas y otras propiedades, para identificar las restricciones que deben ser activas en la siguiente iteración.

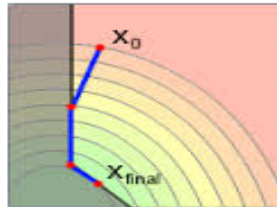
MÉTODOS DE TIPO ACTIVE SET VS. PENALIZACIÓN

Los métodos que identifican (e imponen) las restricciones que deben considerarse activas en cada iteración se conocen, genéricamente, como métodos de tipo **conjunto activo** (active set). Presentan dos ventajas principales frente a los métodos de penalización:

- Imponen las restricciones activas en cada punto.
- Evitan el mal condicionamiento de la matriz hessiana que aparece en los métodos de penalización.



Penalización



Conjunto activo

CONDICIONES NECESARIAS DE MÍNIMO

Se considera ahora el caso general:

$$\begin{aligned} \min f(u_1, \dots, u_N), \text{ con} \\ r_k(u_1, \dots, u_N) = 0, \quad \text{para } k = 1, \dots, m, \\ r_k(u_1, \dots, u_N) \geq 0, \quad \text{para } k = m + 1, \dots, n. \end{aligned}$$

Los mínimos deben verificar el sistema de Lagrange (**primera condición necesaria**) para las restricciones activas, es decir

$$\nabla f(\mathbf{u}) = \sum_{\text{rest. activas}} \lambda_j \nabla r_j(\mathbf{u}).$$

Además, si el mínimo está en la frontera, la función objetivo debe ser no decreciente hacia el interior del espacio de diseño, es decir (**segunda condición necesaria**)

$$\nabla f(\mathbf{u})^\top \mathbf{d} \geq 0 \text{ si } \mathbf{d} \in \mathbb{R}^N \text{ es tal que } \mathbf{d}^\top \nabla r_k > 0 \text{ para las rest. unilaterales activas.}$$

Esta condición resulta ser equivalente (lema de Farkas) a imponer que $\lambda_k \geq 0$ para las restricciones unilaterales activas.

LEMA DE FARKAS

Lema de Farkas: Sea \mathbf{A} una matriz de orden $m \times n$ y \mathbf{b} un vector de \mathbb{R}^n . La ecuación

$$\mathbf{Ax} = \mathbf{b} \text{ con } \mathbf{x} \geq \mathbf{0}$$

tiene solución si, y solamente si, no existe $\mathbf{y} \in \mathbb{R}^m$ tal que

$$\mathbf{A}^\top \mathbf{y} \leq \mathbf{0} \text{ y } \mathbf{b}^\top \mathbf{y} > 0.$$

Aplicando el lema de Farkas con $\mathbf{A} = [\nabla r_1, \dots, \nabla r_p]$, $\mathbf{x} = \lambda$ y $\mathbf{b} = \nabla F$, se obtiene la propiedad enunciada en la transparencia anterior. Es esencial, para este argumento, que los vectores $\nabla r_1, \dots, \nabla r_p$ sean linealmente independientes para que haya una solución, a lo sumo, de la ecuación $\mathbf{Ax} = \mathbf{b}$.

CONDICIONES DE KARUSH-KHUN-TUCKER (KKT)

Resumiendo, los mínimos (del interior y de la frontera) del problema

$$\min f(u_1, \dots, u_N), \text{ con}$$

$$r_k(u_1, \dots, u_N) = 0, \quad \text{para } k = 1, \dots, m,$$

$$r_k(u_1, \dots, u_N) \geq 0, \quad \text{para } k = m + 1, \dots, n,$$

deben verificar (**condiciones KKT**) (implementadas en *fmincon* de MatLab)):

$$\nabla f = \lambda_1 \nabla r_1 + \dots + \lambda_n \nabla r_n.$$

$$r_j = 0 \text{ para } j = 1, \dots, m,$$

$$\lambda_j r_j = 0, \quad \text{y } r_j > 0 \text{ o } \lambda_j \geq 0 \text{ para } j = m + 1, \dots, n.$$

Nótese que **la última condición sólo se aplica a las restricciones unilaterales** e implica que, o bien es $\lambda_j = 0$ y $r_j > 0$ (restricciones unilaterales inactivas) o debe ser $r_j = 0$ y $\lambda_j \geq 0$ (restricciones unilaterales activas).

MÉTODO DE NEWTON Y PROGRAMACIÓN CUADRÁTICA (I)

Dos **formulaciones equivalentes** en cada paso de la iteración, extendiendo sus homólogas para el caso de restricciones unilaterales (las restricciones activas y no activas se denotan con los superíndices a y i , respectivamente):

- **Método de Newton aplicado al sistema de Lagrange:**

$$\begin{aligned} [\partial_{uu}f - \lambda^{a\top} \partial_{uu}r^a]_k(\mathbf{u} - \mathbf{u}_k) - (\lambda^a - \lambda_k^a)^\top [\partial_{ur}r^a]_k &= -[\partial_{uf} - \lambda^{a\top} \partial_{ur}r^a]_k, \\ [\partial_{ur}r^a]_k(\mathbf{u} - \mathbf{u}_k) &= -\mathbf{r}_k, \quad [\partial_{ur}r^i]_k(\mathbf{u} - \mathbf{u}_k) > -\mathbf{r}_k^i, \end{aligned}$$

donde la desigualdad entre vectores se define como desigualdad componente a componente.

- **Programación cuadrática:**

$$\begin{aligned} \min_{\mathbf{u}} f_k + [\partial_{uf} - \lambda^{a\top} \partial_{ur}r^a]_k(\mathbf{u} - \mathbf{u}_k) + \frac{1}{2}(\mathbf{u} - \mathbf{u}_k)^\top [\partial_{uu}f - \lambda^{a\top} \partial_{uu}r^a]_k(\mathbf{u} - \mathbf{u}_k) \\ [\partial_{ur}r^a]_k(\mathbf{u} - \mathbf{u}_k) = -\mathbf{r}_k, \quad [\partial_{ur}r^i]_k(\mathbf{u} - \mathbf{u}_k) > -\mathbf{r}_k^i. \end{aligned}$$

Nótese que las restricciones activas e inactivas pueden cambiar de una iteración a otra.

MÉTODO DE NEWTON-PROGRAMACIÓN CUADRÁTICA (II)

Las dos formulaciones anteriores se implementan como en el caso de restricciones sólo estrictas, pero monitorizando los valores de las restricciones inactivas y los multiplicadores de Lagrange de las restricciones activas unilaterales. Pueden darse tres situaciones en cada iteración:

- 1 $r_k^i > 0$ para toda restricción inactiva y $\lambda_k^a > 0$ para toda restricción activa unilateral. El proceso continúa tal cual.
- 2 $r_k^i < 0$ para alguna restricción inactiva. Se para el proceso iterativo, se localiza el punto \mathbf{u}_0 donde $r_k^i = 0$, se convierte esa restricción en activa, y se continúa el proceso, tomando $\mathbf{u}_{k+1} = \mathbf{u}_0$.
- 3 $\lambda_k^a < 0$ para alguna restricción activa unilateral. Se para el proceso iterativo, se localiza el punto \mathbf{u}_0 donde $\lambda_k^a = 0$, se convierte esa restricción en inactiva, y se continúa el proceso, tomando $\mathbf{u}_{k+1} = \mathbf{u}_0$.

En los casos 2 y 3 de la formulación mediante programación cuadrática, puede sustituirse el proceso (iterativo) necesario para localizar el punto \mathbf{u}_0 (donde $\mathbf{u}_{k+1} = \mathbf{u}_0$ o $\lambda_k^a = 0$) por un **método de penalización** (sobre la aproximación cuadrática), definiendo las funciones de penalización basadas en las restricciones.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 **Multiplicadores de Lagrange y condiciones KKT**
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - **Optimización mediante la herramienta MatLab**
- 4 Cálculo variacional en dimensión infinita
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

HERRAMIENTA *fmincon* DE MATLAB: opciones

- **Trust region reflective.** Sólo restricciones estrictas lineales y restricciones unilaterales del tipo $a_j \leq u_j \leq b_j$: **minimización en caja (box problem)**. Combina Newton (dando o aproximando H) con regiones de confianza.
- **Active set.** Combina BFGS con descenso sobre la formulación mediante programación cuadrática, penalizada para redefinir en cada paso las restricciones unilaterales activas/inactivas.
- **SQP.** Mejora de 'active-set', con dos nuevos ingredientes principales: (i) se obliga a que se verifiquen las restricciones activas en cada paso, utilizando una aproximación cuadrática de estas restricciones combinada con un proceso iterativo, y (ii) se utilizan subrutinas apropiadas de álgebra lineal para mejorar el uso de la memoria. En general, requiere algunas iteraciones más que active set y da resultados más precisos.
- **Interior-point.** Elimina las restricciones unilaterales mediante funciones barrera logarítmicas. Debe darse gradiente; la matriz hessiana puede darse o aproximarse mediante BFGS, LBFGS (**low memory BFGS**, una variante de BFGS que minimiza el uso de memoria mediante una aproximación de baja dimensión de la aproximación BFGS) o diferencias finitas.

HERRAMIENTA *fmincon* DE MATLAB: aplicación 2D (I)

Para minimización en una caja pueden utilizarse cualquiera de los algoritmos.

Para el caso general, en cambio, no puede utilizarse *trust region reflective*, que es el más preciso en principio (aunque requiere calcular la matriz hessiana, que puede ser muy costoso para grandes sistemas); se consideran los tres restantes.

En todos los casos, se utilizarán las opciones que tiene MatLab por defecto.

Las tablas de resultados son similares a las del caso sin restricciones. Se considerarán los siguientes algoritmos:

- *Trust region reflective*, dando el vector gradiente y la matriz hessiana (TRR+H+G) o dando sólo el vector gradiente y aproximando la matriz hessiana por diferencias finitas (TRR+G).
- *SQP*, dando el vector gradiente (SQP+G) o aproximándolo por diferencias finitas (SQP).
- *Active set*, dando el vector gradiente (AS+G) o aproximándolo por diferencias finitas (AS).
- *Interior point*, aproximando la matriz hessiana mediante BFGS (IP+BFGS) o mediante diferencias finitas (IP). No se utilizará LBGS, que en 2D da resultados prácticamente idénticos a BFGS.

HERRAMIENTA *fmincon* DE MATLAB: aplicación 2D (II)

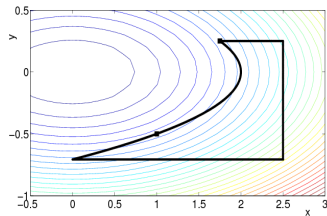
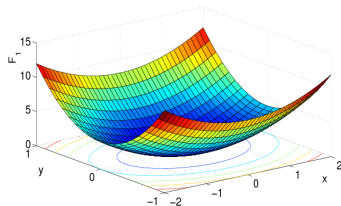
Deben darse la función objetivo, los límites superior e inferior de las variables (igual que en *fminunc*) y las restricciones (cuidado, MatLab formula las restricciones unilaterales con \leq , en vez de \geq , como hemos hecho nosotros):

- **Estrictas lineales:** $\mathbf{A}\mathbf{u} = \mathbf{b}$ se dan en 'Linear equalities' dando la matriz \mathbf{A} y el vector \mathbf{b} como 'A' y 'b', respectivamente.
- **Unilaterales lineales:** $\mathbf{A}\mathbf{u} \leq \mathbf{b}$ se dan en 'Linear inequalities' dando la matriz \mathbf{A} y el vector \mathbf{b} como 'Aeq' y 'beq', respectivamente.
- **No lineales:** $\mathbf{r}^u(u_1, u_2) \leq \mathbf{0}$ y $\mathbf{r}^e(u_1, u_2) = \mathbf{0}$. Deben darse \mathbf{r}^u y \mathbf{r}^e y sus gradientes. Por ejemplo, para la restricción no lineal de la transparencia siguiente, donde no hay restricciones estrictas, puede indicarse '@constraint' y dar la función en un fichero 'constraint.m' (que contine la restricción y su gradiente):

```
function [c,ceq,DC,DCEq] = constraint(x)
c = 2 - x(1) - 4 * x(2)^2;    ceq = [];
if nargout>2
DC = [-1; -4 * x(2)];    DCEq = [];
end
```

APLICACIÓN: FUNCIÓN CONVEXA CON RESTRICCIONES (I)

- Función objetivo: $F_1(x, y) \equiv x^2 + 8y^2$.
- Restricciones: $2 - 4y^2 \leq x \leq 2.5$, $-1/\sqrt{2} \leq y \leq 1/4$; la restricción no lineal impide utilizar el método TRR.
- Dos mínimos, ambos en la frontera: global en $(x, y) = (1, -1/2)$; local en $(x, y) = (7/4, 1/4)$.



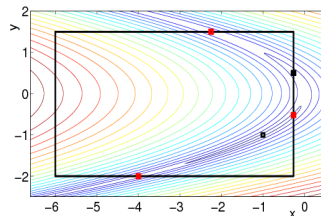
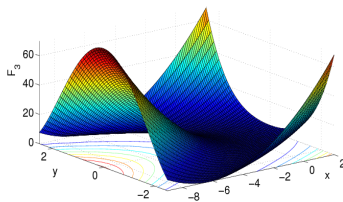
APLICACIÓN: FUNCIÓN CONVEXA CON RESTRICCIONES (II)

init. point	(2,0.25)			(2,0.2)			(2,-0.25)		
algorithm	iter.	eval.	error	iter.	eval.	error	iter.	eval.	error
SQP+G	2	3	10^{-16} (L)	3	4	10^{-16}	4	6	$8 \cdot 10^{-7}$
SQP	2	9	10^{-16} (L)	3	12	10^{-16}	4	16	$5 \cdot 10^{-6}$
AS+G	1	3	10^{-16} (L)	2	5	10^{-16}	4	10	$2 \cdot 10^{-7}$
AS	1	6	10^{-16} (L)	2	9	$4 \cdot 10^{-11}$	4	16	$2 \cdot 10^{-7}$
IP+G	6	22	$3 \cdot 10^{-7}$ (G)	9	10	$2 \cdot 10^{-6}$	7	9	$2 \cdot 10^{-9}$
IP	16	83	$6 \cdot 10^{-7}$ (G)	9	30	$2 \cdot 10^{-6}$	7	25	$4 \cdot 10^{-9}$

- Se da el gradiente de las restricciones en todos los casos; el número de evaluaciones aumenta si se calcula por diferencias finitas.
- Mejores prestaciones: en el orden que se dan. IP funciona mucho peor que los demás métodos.
- Las condiciones iniciales (2,0.2) y (2,-0.25) están en los dominios de atracción del mínimo local y global, respectivamente;
- Sorprendentemente, la condición (2,0.25) conduce al mínimo local (L) o al global (G). La razón de que IP converja al mínimo global, que está más alejado, es que esta condición inicial está en la frontera, y la función barrera que utiliza IP penaliza mucho la función objetivo en la primera iteración.

APLICACIÓN: PROBLEMA EN UNA CAJA PARA LA FUNCIÓN BANANA (I)

- Función objetivo: $F_3(x, y) \equiv (x + y^2)^2 + (1 + y)^2/100$.
- Restricciones: $-6 \leq x \leq -53/200$, $-2 \leq y \leq 3/2$.
- Dos mínimos (rectángulos negros en la figura): global en $(x, y) = (-1, -1)$; local en $(x, y) = (-53/200, 1/2)$.
- La restricción a la frontera tiene dos mínimos locales adicionales (rectángulos rojos), pero éstos no son mínimos del problema de optimización porque los multiplicadores de Lagrange son negativos.



APLICACIÓN: PROBLEMA EN UNA CAJA PARA LA FUNCIÓN BANANA (II)

- **Condiciones iniciales:** $(-6,1)$ y $(-6,-1)$ convergen al mínimo local y global, respectivamente, y $(-6,0.5)$ converge a uno de los dos, según el caso (se indica con las iniciales L y G entre paréntesis).

pto. inic.	(-6,1)			(-6,0.5)			(-6,-1)		
	iter.	eval.	error	iter.	eval.	error(min.)	iter.	eval.	error
TRR+H+G	16	-	$4 \cdot 10^{-7}$	16	-	$3 \cdot 10^{-6}$ (L)	17	-	$6 \cdot 10^{-6}$
TRR+G	16	-	$4 \cdot 10^{-7}$	16	-	$3 \cdot 10^{-6}$ (L)	17	-	$6 \cdot 10^{-6}$
SQP+G	7	8	$2 \cdot 10^{-8}$	22	27	$2 \cdot 10^{-6}$ (G)	26	32	$4 \cdot 10^{-7}$
SQP	7	24	$2 \cdot 10^{-8}$	22	73	10^{-5} (G)	26	86	10^{-5}
AS+G	5	11	$3 \cdot 10^{-4}$	19	41	$4 \cdot 10^{-3}$ (G)	21	52	$7 \cdot 10^{-3}$
AS	5	18	$3 \cdot 10^{-4}$	17	60	$4 \cdot 10^{-3}$ (G)	21	75	$7 \cdot 10^{-3}$
IP+G	28	31	$3 \cdot 10^{-6}$	24	25	10^{-5} (L)	26	28	10^{-5}
IP	28	89	$3 \cdot 10^{-6}$	24	75	$3 \cdot 10^{-6}$ (L)	26	82	$5 \cdot 10^{-6}$

- Se da el gradiente de las restricciones en todos los casos; el número de evaluaciones aumenta si se calcula por diferencias finitas.
- Mejores prestaciones: en el orden que se dan.

OPTIMIZACIÓN CON RESTRICCIONES: CONCLUSIONES

Una vez desechados métodos menos eficaces, como los métodos de penalización, o los basados en quasi-Newton-DFP y descenso más empinado:

- Los métodos de tipo **conjunto activo**, que hacen uso de multiplicadores de Lagrange y condiciones KKT son el estado del arte a día de hoy. Estos métodos están implementados en los algoritmos *DFP active-set* de MatLab.
- Ambos algoritmos dan resultados parecidos en 2D. Para dimensión alta, en cambio, DFP es preferible.
- Ambos algoritmos dan buenos resultados incluso en situaciones muy exigentes.
- Por sencillez, la aplicación se ha hecho en 2D, pero en **dimensión mayor** se obtienen conclusiones parecidas: añadiendo solamente 'dimensiones pasivas' aumenta el número de evaluaciones; si, en cambio, las nuevas dimensiones contienen complejidad, el número de iteraciones también aumenta.

INTRODUCCIÓN AL CÁLCULO DE VARIACIONES (I)

- En muchos problemas de interés práctico, la función objetivo es un funcional y los parámetros de diseño son funciones (en vez de vectores).
- El valor óptimo del funcional se obtiene eligiendo una o varias **funciones de diseño** (en vez de parámetros de diseño).
- Estos problemas son **generalizaciones a dimensión infinita** de los problemas considerados hasta ahora. Las funciones son vectores de dimensión infinita (o muy grande si se discretizan).
- Ejemplo 1: optimizar una **trayectoria de un satélite** (funciones del tiempo) minimizando el tiempo de vuelo (integral a lo largo de la trayectoria).
- Ejemplo 2: obtener la **forma de un ala** (funciones de dos variables espaciales) minimizando la resistencia aerodinámica (una integral a lo largo del ala).

INTRODUCCIÓN AL CÁLCULO DE VARIACIONES (II)

- Podrá haber **una o varias funciones de diseño**.
- Las funciones de diseño podrán depender de **una o varias variables independientes**.
- Sólomente se consideran **funcionales definidos a través de integrales** extendidas a dominios acotados e integrales sobre las fronteras de estos dominios, con integrandos dependientes de las funciones de diseño y de sus derivadas.
- La generalización de la condición necesaria de gradiente de la función objetivo nulo consiste en una o varias **ecuaciones diferenciales**, ordinarias o en derivadas parciales, con **condiciones iniciales o de contorno**.

INTRODUCCIÓN AL CÁLCULO DE VARIACIONES (III)

- Se deducen **condiciones necesarias de extremo**, que se llamarán siempre, en lo que sigue, **ecuaciones Euler**; en algunos textos, se emplea un segundo nombre (Lagrange, Poisson, Ostrogradski, etc).
- También aparecen condiciones de contorno, de dos modos:
 - Como parte del proceso de optimización (**condiciones de contorno naturales**).
 - Porque se imponen a las funciones admisibles; estas condiciones de contorno estarán sujetas a limitaciones.

INTRODUCCIÓN AL CÁLCULO DE VARIACIONES (IV)

Lo anterior corresponde al **cálculo de variaciones 'elemental'**, que puede verse como un problema 'limpio' de optimización sin restricciones. Pero hay **otras situaciones de interés práctico**, que también se tratarán brevemente al final de este apartado:

- Las variables de diseño pueden incluir **funciones y escalares**. En este caso, aparece una condición adicional por cada parámetro de diseño escalar.
- Además, pueden imponerse **restricciones**, estrictas o unilaterales.
- Las restricciones estrictas dan lugar a **ecuaciones de Euler con multiplicadores de Langange**.
- Si las restricciones estrictas incluyen ecuaciones diferenciales, conviene formular las ecuaciones de Euler mediante **operadores adjuntos**.

INTRODUCCIÓN AL CÁLCULO DE VARIACIONES (V)

- En lo que sigue, solamente se consideran algunos casos generales para ilustrar los **métodos básicos**. La variedad de problemas (número de variables independientes, número de funciones a determinar, orden de las derivadas que aparecen en el funcional, etc.) es demasiado grande para intentar un estudio exhaustivo en este curso.
- Se considera, en primer lugar, **un caso general** en que la función objetivo depende de una sola función de diseño dependiente de dos variables independientes (2D) y de sus derivadas primeras. La restricción a 1-D y la generalización a dimensión mayor son sencillas.
- El resto de los casos (dependencia de derivadas de orden superior, varias funciones de diseño, etc.) se consideran sólo para **ejemplos representativos concretos**.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 Multiplicadores de Lagrange y condiciones KKT
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 **Cálculo variacional en dimensión infinita**
 - **Cálculo de variaciones 'elemental'**
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

LEMA FUNDAMENTAL DEL CÁLCULO DE VARIACIONES

Éste es el resultado matemático que nos permitirá deducir condiciones de extremo:

- Sea $\Omega \in \mathbb{R}^n$, con $n \geq 1$ un dominio (conjunto abierto y conexo) de frontera suave a trozos y sea $f : \Omega \rightarrow \mathbb{R}$ una función continua. Si

$$\int_{\Omega} f \tilde{u} d\mathbf{x} = 0, \text{ con } d\mathbf{x} = dx_1 \dots dx_n,$$

para toda función $\tilde{u} \in C^\infty(\Omega)$, que sea idénticamente nula en un entorno de la frontera (esta condición define a **funciones de soporte compacto**), entonces la función $f = 0$ en Ω .

- Nótese que una función con soporte compacto es tal que $\tilde{u} = \partial_\nu \tilde{u} = \partial_{\nu\nu} \tilde{u} = \dots = 0$ y, por tanto, verifica cualquier condición de contorno homogénea que involucre a la función y a su derivadas, de cualquier orden, respecto de la normal unitaria exterior ν , tales como $\tilde{u} = 0$ (Dirichlet) o $\partial_\nu \tilde{u} = 0$ (Neumann).

FORMULACIÓN GENERAL (I)

- Considérese el problema de **seleccionar la función** $u = u(x_1, \dots, x_n)$ **para minimizar el funcional**

$$F(u) = \int_{\Omega} f(\mathbf{x}, u, Du, D^2u, \dots, D^m u) d\mathbf{x} \\ + \int_{\partial\Omega} g(\mathbf{x}, u, \partial_{\nu} u, \dots, \partial_{\nu \dots \nu}^k u) dA, \quad \text{con } k < m,$$

donde f y g son funciones suaves, $\Omega \subset \mathbb{R}^n$ es un dominio con frontera (que se denota como $\partial\Omega$) suave a trozos, $D^j u$ denota la diferencial de orden j (conjunto de las derivadas parciales de ese orden), y $\partial_{\nu \dots \nu}^j u$ denota la derivada de orden j de u respecto de la normal unitaria exterior.

- Nótese que, en la frontera, el funcional involucra derivadas de, a lo sumo, un orden menor que el funcional en el interior del dominio.

FORMULACIÓN GENERAL (II)

- En principio, las **funciones admisibles** (sobre las que se minimiza el funcional) son derivables con continuidad m veces (para que el funcional esté bien definido). Sin embargo, se exigirá que **las funciones admisibles sean derivables $2m$ veces** porque, como se verá, ese es el orden de las derivadas que aparecerán en las ecuaciones de Euler.
- Si no se impone ninguna condición adicional a las funciones admisibles, el proceso de minimización producirá m condiciones de contorno, que se denominarán **naturales**.
- También pueden imponerse hasta m condiciones de contorno a las funciones admisibles, que podrán involucrar a la propia función y a sus derivadas respecto de ν , hasta orden $m - 1$, a lo sumo.

CÁLCULO DE VARIACIONES EN 2D (I)

- Considérese el problema de **seleccionar la función u para minimizar el funcional**

$$F(u) = \int_{\Omega} f(\partial_x u, \partial_y u, u, x, y) \, dx dy + \int_{\partial\Omega} g(u, s) \, ds,$$

donde $f = f(p, q, u, x, y)$ y $g = g(u, s)$ son funciones suaves, $\Omega \subset \mathbb{R}^2$ es un dominio con frontera $(\partial\Omega)$ suave a trozos, y s es un parámetro longitud de arco a lo largo de la frontera.

- Nótese que, en la frontera, el funcional involucra derivadas de un orden menor que el funcional en el interior del dominio; es decir, no involucra a las derivadas en este caso.
- **Enfoque variacional:** Si $u(x, y)$ minimiza el funcional, sustituyendo $u = u(x, y) + \alpha \tilde{u}(x, y)$ en el funcional, la expresión resultante debe tener un mínimo en $\alpha = 0$. Por tanto, la derivada de esa expresión respecto de α debe ser nula en $\alpha = 0$.

CÁLCULO DE VARIACIONES EN 2D (II)

- Se obtiene así la ecuación

$$\int_{\Omega} [\partial_x \tilde{u} \partial_p f + \partial_y \tilde{u} \partial_q f + \tilde{u} \partial_u f] dx dy + \int_{\partial\Omega} \tilde{u} \partial_u g ds = 0,$$

que debe verificarse para toda función admisible (derivable con continuidad).

- Integrando por partes en la expresión anterior, se obtiene

$$\int_{\Omega} [-\partial_x (\partial_p f) - \partial_y (\partial_q f) + \partial_u f] \tilde{u} dx dy + \int_{\partial\Omega} \tilde{u} [\partial_u g + \boldsymbol{\nu} \cdot (\partial_p f, \partial_q f)] ds = 0,$$

donde $\boldsymbol{\nu}$ es la normal unitaria exterior a la frontera del dominio. Y esta ecuación debe verificarse para toda función admisible \tilde{u} .

ECUACIÓN DE EULER

- Considerando sólo funciones \tilde{u} con soporte compacto (para las que $\tilde{u} = 0$ en $\partial\Omega$), y utilizando el lema fundamental del cálculo de variaciones, la condición

$$\int_{\Omega} [-\partial_x(\partial_p f) - \partial_y(\partial_q f) + \partial_u f] \tilde{u} \, dx dy + \int_{\partial\Omega} [\partial_u g + \boldsymbol{\nu} \cdot (\partial_p f, \partial_q f)] \tilde{u} \, ds = 0,$$

implica que (condición necesaria de extremo)

$$-\partial_x(\partial_p f) - \partial_y(\partial_q f) + \partial_u f = 0 \text{ en } \Omega.$$

- Ésta es, en general, una ecuación diferencial, en derivadas parciales, que recibe el nombre de **ecuación de Euler**.

CONDICIONES DE CONTORNO NATURALES

- Por tanto (como la primera integral es nula), la condición original se reduce a

$$\int_{\partial\Omega} [\partial_u g + \boldsymbol{\nu} \cdot (\partial_p f, \partial_q f)] \tilde{u} \, ds = 0$$

para toda función \tilde{u} continua sobre la frontera. Esto implica que

$$\partial_u g + \boldsymbol{\nu} \cdot (\partial_p f, \partial_q f) = 0 \text{ en } \partial\Omega.$$

- Esta condición de contorno (que no se impuso a las funciones admisibles) aparece al minimizar. Recibe el nombre de **condición de contorno natural**.

OTRAS CONDICIONES DE CONTORNO

- Si, en cambio, se hubiera impuesto una condición del tipo $u = h(s)$ a las funciones admisibles, ésta sería precisamente la condición de contorno a imponer a la ecuación de Euler. Nótese que la condición de contorno a imponer a las funciones admisibles debe involucrar derivadas de un orden inferior a las que aparecen en la integral doble, es decir, debe depender sólo de u en este caso.
- Con esa condición de contorno, todo es consistente, ya que:
 - La integral de línea que aparece en el funcional es constante, y no juega ningún papel en la minimización.
 - \tilde{u} debe ser idénticamente nula en la frontera (para que $u + \alpha\tilde{u}$ verifique la condición de contorno), de modo que la integral de línea que aparece al integrar por partes es nula.

ECUACIONES DE ORDEN SUPERIOR: UN EJEMPLO (I)

- Se procede igual, pero hay que integrar por partes varias veces.
- Por ejemplo, para minimizar

$$\int_{\Omega} [(\Delta u)^2 + |\nabla u|^2 - u + u^2] dx dy$$

se sustituye $u = u + \alpha \tilde{u}$, y se impone que la derivada respecto de α en $\alpha = 0$ sea nula, para obtener

$$\int_{\Omega} [2\Delta u \Delta \tilde{u} + 2\nabla u \cdot \nabla \tilde{u} - \tilde{u} + 2u\tilde{u}] dx dy = 0.$$

- O, integrando por partes dos veces y dividiendo por 2,

$$\int_{\Omega} [\Delta^2 u - \Delta u + u - 1/2] \tilde{u} dx dy + \int_{\partial\Omega} [\Delta u \partial_{\nu} \tilde{u} + (\partial_{\nu} u - \partial_{\nu} \Delta u) \tilde{u}] ds = 0,$$

para toda función admisible \tilde{u} (con derivadas parciales segundas continuas); ∂_{ν} es la derivada respecto de la normal unitaria exterior.

ECUACIONES DE ORDEN SUPERIOR: UN EJEMPLO (II)

- Imponiendo la condición anterior para toda función \tilde{u} de soporte compacto (para las cuales $\tilde{u} = \partial_\nu \tilde{u} = 0$ en $\partial\Omega$), se obtiene la **ecuación de Euler**

$$\Delta^2 u - \Delta u + u - 1/2 = 0 \text{ en } \Omega.$$

- Por tanto, la integral doble es nula y la condición queda

$$\int_{\partial\Omega} [\Delta u \partial_\nu \tilde{u} + (\partial_\nu u - \partial_\nu \Delta u) \tilde{u}] ds = 0$$

para toda función \tilde{u} derivable dos veces; \tilde{u} y $\partial_\nu \tilde{u}$ pueden tomar cualquier valor sobre la frontera, independientes una de otra. Por tanto, los coeficientes de \tilde{u} y $\partial_\nu \tilde{u}$ deben ser nulos, y se obtienen las siguientes **condiciones de contorno naturales**:

$$\Delta u = 0, \quad \partial_\nu u - \partial_\nu (\Delta u) = 0 \text{ en } \partial\Omega.$$

ECUACIONES DE ORDEN SUPERIOR: UN EJEMPLO (III)

- Las condiciones de contorno naturales aparecen sin haber impuesto ninguna condición de contorno a las funciones admisibles.
- Si, en cambio, se impusiese **una condición de contorno** del tipo (la condición de contorno debe involucrar solamente a u y a $\partial_\nu u$)

$$g(u, \partial_\nu u) = 0 \quad \text{en } \partial\Omega$$

a las funciones admisibles, se tendría

$$\tilde{u} \partial_u g + \partial_\nu \tilde{u} \partial_p g = 0$$

(¡ojo! como la condición de contorno anterior es no lineal, las perturbaciones que la verifiquen no pueden ser de la forma $u + \alpha \tilde{u}$; deben considerarse perturbaciones más generales, $u = U(x, y, \alpha)$, de modo que en la expresión anterior $u = U(x, y, 0)$ y $\tilde{u} = \partial_\alpha U(x, y, 0)$). Sustituyendo la expresión en la integral de línea, se obtendría la siguiente condición de contorno natural (además de la condición anterior)

$$\partial_\nu (\Delta u) \partial_p g + (\Delta u + u) \partial_u g = 0 \quad \text{en } \partial\Omega.$$

ECUACIONES DE ORDEN SUPERIOR: UN EJEMPLO (IV)

- Si se impusiesen **dos condiciones de contorno independientes** (que determinen unívocamente a u y a $\partial_\nu u$ en cada punto de la frontera) del tipo

$$g_1(u, \partial_\nu u) = 0, \quad g_2(u, \partial_\nu u) = 0 \quad \text{en } \partial\Omega,$$

tales condiciones se impondrían directamente a las ecuaciones de Euler y no aparecería ninguna condición de contorno natural.

- Nótese que las condiciones de contorno podrían escribirse como $u = h_1(s)$ y $\partial_\nu u = h_2(s)$, y todo sería consistente porque:
 - La integral de línea que aparece en el funcional es constante y puede ignorarse al minimizar.
 - Las funciones \tilde{u} serían tales que $\tilde{u} = \partial_\nu \tilde{u} = 0$ en $\partial\Omega$ (para que $u + \alpha \tilde{u}$ verifique las condiciones de contorno) y, por tanto, la integral de línea que aparece al integrar por partes sería idénticamente nula.

ECUACIONES DE ORDEN SUPERIOR: CONCLUSIONES (I)

Aunque se ha considerado sólo un ejemplo, las siguientes conclusiones tienen **validez general**, en casos genéricos en que:

- El funcional depende de derivadas parciales de hasta orden dos de la variable de diseño.
- Considerando sólo las derivadas segundas, se tiene un funcional definido positivo. Esta condición es necesaria para que el funcional esté acotado inferiormente y pueda tener un mínimo global.

La ecuación de Euler que se obtiene (al aplicar el método variacional) **es una ecuación elíptica de cuarto orden** y, por tanto, requiere **dos condiciones de contorno independientes**.

ECUACIONES DE ORDEN SUPERIOR: CONCLUSIONES (II)

Para las dos condiciones de contorno de la ecuación de Euler, se pueden dar tres situaciones:

- **Si no se imponen condiciones de contorno** a las funciones admisibles, el proceso de minimización produce, precisamente, dos condiciones de contorno naturales independientes.
- **Si se impone una condición de contorno** a las funciones admisibles (¡ojo!, ésta no puede involucrar derivadas de orden superior al primero), el proceso de minimización produce una condición de contorno natural. Esa condición adicional se impone, junto con la condición impuesta de entrada.
- **Si se imponen dos condiciones de contorno** independientes a las funciones admisibles, éstas son las que deben imponerse a la ecuación de Euler (no hay condiciones de contorno naturales a adicionales).

VARIAS VARIABLES DEPENDIENTES: UN EJEMPLO (I)

Nuevamente, se considera un ejemplo: minimizar el funcional

$$\int_{\Omega} [|\nabla u|^2 + |\nabla v|^2 + 2uv] \, dx dy.$$

Para ello, se sustituyen $u = u + \alpha \tilde{u}$ y $v = v + \alpha \tilde{v}$ en el funcional. En la expresión resultante, se iguala a cero la derivada respecto de α en $\alpha = 0$ y se integra por partes en la expresión resultante, para obtener

$$\begin{aligned} \int_{\Omega} [\nabla u \cdot \nabla \tilde{u} + \nabla v \cdot \nabla \tilde{v} + \tilde{u}v + u\tilde{v}] \, dx dy \equiv \\ \int_{\Omega} [(v - \Delta u)\tilde{u} + (u - \Delta v)\tilde{v}] \, dx dy + \int_{\partial\Omega} (\tilde{u}\partial_{\nu} u + \tilde{v}\partial_{\nu} v) \, ds = 0 \end{aligned}$$

para todo par de funciones admisibles \tilde{u} y \tilde{v} .

Procediendo como en los casos anteriores, las ecuaciones de Euler son

$$\Delta u - v = \Delta v - u = 0 \text{ en } \Omega.$$

VARIAS VARIABLES DEPENDIENTES: UN EJEMPLO (II)

Se dan tres situaciones para las condiciones de contorno:

- Si no se impone ninguna condición de contorno a las funciones admisibles, las condiciones de contorno naturales son

$$\partial_\nu u = \partial_\nu v = 0 \text{ en } \partial\Omega.$$

- Si se impone la condición de contorno

$$g(u, v, s) = 0 \text{ en } \partial\Omega$$

(no puede involucrar derivadas de u y v), las perturbaciones de las funciones admisibles deben verificar $(\partial_u g)\tilde{u} + (\partial_v g)\tilde{v} = 0$, y se obtiene la condición natural (que se añade a la anterior)

$$[\partial_\nu g(s, u, v)]\partial_\nu u - [\partial_u g(s, u, v)]\partial_\nu v = 0 \text{ en } \partial\Omega.$$

- Si se imponen dos condiciones de contorno independientes a las funciones admisibles, éstas son las que deben imponerse a las ecuaciones de Euler.

EJEMPLOS

Para resolver directamente, aplicando el método visto anteriormente

- $F = \int_0^1 [|u'(x)|^2/2 + xu] dx$, (i) sin condiciones adicionales, o (ii) con la condición de contorno $u = 1$ en $x = 1$.
- $F = \int_{\Omega} [|\nabla u|^2/2 + (x^2 + y^2)u] dxdy$, (i) sin condiciones adicionales, o (ii) con la condición de contorno $u = 1$ en $\partial\Omega$.
- $F = \int_{\Omega} [|\nabla u|^2/2 + (x^2 + y^2)u] dxdy + \int_{\partial\Omega} u^2 ds$, (i) sin condiciones adicionales, o (ii) con la condición de contorno $u = 1$ en $\partial\Omega$.
- $F = \int_{\Omega} |\nabla u|^2 dxdy / \int_{\Omega} u^2 dxdy$, (i) sin condiciones adicionales, o (ii) con la condición de contorno $u = 0$ en $\partial\Omega$.
- $F = \int_{\Omega} [|\Delta u|^2/2 + (x^2 + y^2)u] dxdy$, (i) sin condiciones adicionales, o (ii) con la condición de contorno $u = 1$ en $\partial\Omega$.

En todos los casos, Ω es el círculo de radio unidad, con centro en el origen.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 Multiplicadores de Lagrange y condiciones KKT
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 **Cálculo variacional en dimensión infinita**
 - Cálculo de variaciones 'elemental'
 - **Variables de diseño escalares adicionales**
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - Optimización vs. discretización

VARIABLES DE DISEÑO ESCALARES ADICIONALES: UN EJEMPLO (I)

Por sencillez, se minimiza la función objetivo particular

$$F(u) = \int_{\Omega} [(1 + u^2)|\nabla u|^2 - Pu + P^2 + u^4] dx dy,$$

para calcular la función u y el escalar P . Se consideran 'curvas' del espacio de diseño, se sustituye $u = u + \alpha \tilde{u}$, $P = P + \alpha \tilde{P}$, en el funcional, y se impone que la derivada respecto de α se anule en $\alpha = 0$. Integrando por partes, en la ecuación resultante se obtiene

$$\begin{aligned} & \int_{\Omega} [-2\nabla \cdot (1 + u^2)\nabla u] + 2u|\nabla u|^2 - P + 4u^3] \tilde{u} dx dy \\ & + \int_{\Omega} (2P - u)\tilde{P} dx dy + 2 \int_{\partial\Omega} (1 + u^2)(\partial_{\nu} u)\tilde{u} ds = 0 \end{aligned}$$

para toda función \tilde{u} y todo escalar \tilde{P} .

VARIABLES DE DISEÑO ESCALARES ADICIONALES: UN EJEMPLO (II)

- La condición anterior implica que

$$-2\nabla \cdot (1 + u^2)\nabla u + 2u|\nabla u|^2 - P + 4u^3 = 0 \text{ en } \Omega, \quad \partial_\nu u = 0 \text{ en } \partial\Omega,$$

$$\int_{\Omega} (2P - u) \, dx dy = 0.$$

- Se tiene así una incógnita adicional, P , y una ecuación escalar adicional.
- Como hasta ahora, ha aparecido la **condición de contorno natural** $\partial_\nu u = 0$ por no haber impuesto ninguna condición de contorno a las funciones admisibles. Si se hubiese impuesto

$$u = g(x, y) \text{ en } \partial\Omega,$$

que involucra sólo a la función u , como en los casos anteriores, sería esta condición la relevante para integrar las ecuaciones de Euler.

VARIABLES DE DISEÑO ESCALARES ADICIONALES: UN EJEMPLO (III)

Procediendo del mismo modo, en general, si se tienen m funciones y n escalares, se obtiene un sistema acoplado, con:

- m ecuaciones de Euler con sus condiciones de contorno, una por cada función,
- n ecuaciones escalares, una por cada variable escalar.

Por tanto, el sistema resultante tiene 'tantas ecuaciones como incógnitas'.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 Multiplicadores de Lagrange y condiciones KKT
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 **Cálculo variacional en dimensión infinita**
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - **Restricciones estrictas escalares: multiplicadores de Lagrange**
 - Optimización vs. discretización

RESTRICCIONES ERICTAS ESCALARES (I)

Por sencillez, se minimiza la función objetivo

$$F(u) = \int_{\Omega} [(1 + u^2)|\nabla u|^2 - u + u^4] dx dy, \quad \text{con} \quad \int_{\Omega} (u^2 - u) dx dy = 0,$$

que está sujeta a una restricción escalar adicional. El enfoque variacional, haciendo $u = U(x, y, \alpha)$ ¹ (con $U(x, y, 0) = u$ = el minimizante que se está buscando) e integrando por partes, proporciona la ecuación

$$\begin{aligned} \int_{\Omega} \{-2\nabla \cdot [(1 + u^2)\nabla u] + 2u|\nabla u|^2 - 1 + 4u^3\} \tilde{u} dx dy \\ + 2 \int_{\partial\Omega} [(1 + u^2)\partial_{\nu} u] \tilde{u} ds = 0 \end{aligned}$$

para toda función \tilde{u} que verifique

$$\int_{\Omega} (2u - 1)\tilde{u} dx dy = 0.$$

¹ Como en otro caso ya considerado, no vale hacer $U = u + \alpha \tilde{u}$ porque la restricción es no lineal

RESTRICCIONES ERICTAS ESCALARES (II)

Considerando primero funciones \tilde{u} con soporte compacto, se tiene

$$\int_{\Omega} H_1 \, dx dy \equiv \int_{\Omega} \{-2\nabla \cdot [(1 + u^2)\nabla u] + 2u|\nabla u|^2 - 1 + 4u^3\} \tilde{u} \, dx dy = 0$$

para toda función \tilde{u} que verifique

$$\int_{\Omega} H_2 \, dx dy \equiv \int_{\Omega} (2u - 1) \tilde{u} \, dx dy = 0.$$

Con el producto escalar de L_2 ($\langle u_1, u_2 \rangle = \int_{\Omega} u_1 u_2$), esta condición puede expresarse diciendo que la función H_1 debe ser ortogonal a todas las funciones que sean ortogonales a H_2 , lo cual implica que H_1 tiene que ser proporcional a H_2 , es decir,

$$-2\nabla \cdot [(1 + u^2)\nabla u] + 2u|\nabla u|^2 - 1 + 4u^3 = \lambda(2u - 1) \text{ en } \delta\Omega.$$

Razonando como en los casos anteriores, se tiene, además, la condición de contorno natural $\partial_{\nu} u = 0$ en $\partial\Omega$.

RESTRICCIONES ERICTAS ESCALARES (III)

Reuniendo lo anterior, se ha deducido la **ecuación de Euler**

$$-2\nabla \cdot [(1 + u^2)\nabla u] + 2u|\nabla u|^2 - 1 + 3u^3 = \lambda(2u - 1) \text{ en } \delta\Omega,$$

que contiene el **multiplicador de Lagrange** λ , y la condición de contorno natural

$$\partial_\nu u = 0 \text{ en } \partial\Omega.$$

Se tienen, además, la restricción

$$\int_{\Omega} (u^2 - u) \, dx dy = 0.$$

Es decir, se tiene una incógnita adicional, λ , y una restricción adicional a la ecuación de Euler.

Si hubiera m restricciones estrictas, se tendrían m multiplicadores de Lagrange.

EJEMPLO EN 1D

Para el funcional objetivo

$$F_c(u) = \int_0^1 [(1+x^2)(du/dx)^2/10 + u^2(du/dx)/10 + 8u^4 - xu^3 + P^2 - Px^2u] dx,$$

dependiente de la función u tal que $u(\pm 1) = 0$ y $\int_{-1}^1 u dx = 0$, y del escalar P :

- Discretice el funcional en una malla equiespaciada, y resuelva el problema discreto resultante la herramienta MatLab, utilizando al menos dos algoritmos distintos.
- Deduzca la formulación continua, en términos de una ecuación de Euler restringida, y analice lo bien que los óptimos calculados en el punto anterior verifican esta formulación.

Outline

- 1 Algunas cuestiones preliminares
- 2 Métodos de penalización y funciones barrera
- 3 Multiplicadores de Lagrange y condiciones KKT
 - Multiplicadores de Lagrange
 - Condiciones de Karush-Khun-Tucker
 - Optimización mediante la herramienta MatLab
- 4 **Cálculo variacional en dimensión infinita**
 - Cálculo de variaciones 'elemental'
 - Variables de diseño escalares adicionales
 - Restricciones estrictas escalares: multiplicadores de Lagrange
 - **Optimización vs. discretización**

OPTIMIZACIÓN VS. DISCRETIZACIÓN (I)

- En muchos problemas de diseño, la función objetivo es, en realidad, un funcional y los parámetros de diseño son funciones.
- El cálculo numérico de un diseño óptimo requiere discretizar el problema, pero la discretización puede hacerse antes o después de la optimización.
- Si primero se optimiza, se obtienen unas ecuaciones de Euler o ecuaciones adjuntas, como las obtenidas en la segunda parte de este tema. Estas ecuaciones son, en general, diferenciales o integro-diferenciales. Tales ecuaciones deben discretizarse para su tratamiento numérico.
- Si se comienza discretizando el problema, se obtiene un problema de optimización en dimensión finita, como los tratados en la primera parte de este tema.

OPTIMIZACIÓN VS. DISCRETIZACIÓN (II)

- El modo más sensato de proceder es, por tanto, primero discretizar las ecuaciones asociadas a las distintas disciplinas técnicas y luego optimizar el problema de dimensión finita resultante.
- Esto permite un diseño modular de las distintas disciplinas técnicas.
- Las ecuaciones de cada disciplina técnica dependen, en general, sólo de una parte de los parámetros/funciones de diseño.
- Cada disciplina técnica puede tratarse por separado en lo relativo a la consistencia y estabilidad numéricas de cada discretización.
- Todo ello favorece la robustez y flexibilidad de los métodos.